APPENDIX B

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 072 679 A3

(12) EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
31.07.2002 Bulletin 2002/31

(43) Date of publication A2:
31.01.2001 Bulletin 2001/05

(21) Application number: **00306128.0**

(22) Date of filing: **19.07.2000**

(51) Int Cl.[7]: **C12N 15/11**, C12P 19/34, C12Q 1/68, C07H 21/00 // C12N9/12

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **20.07.1999 US 358141**

(71) Applicant: **Agilent Technologies, Inc. (a Delaware corporation)**
Palo Alto, CA 94303 (US)

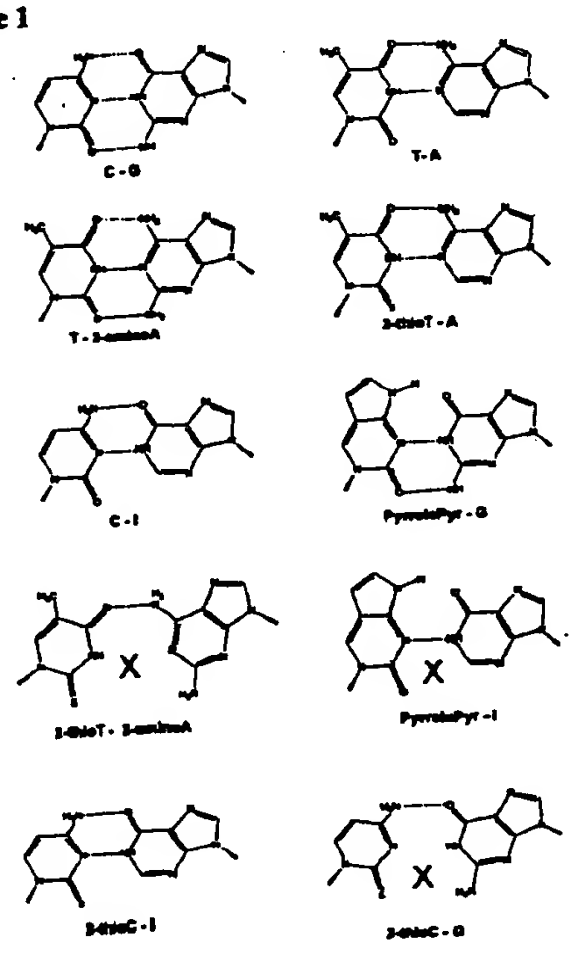
(72) Inventor: **Sampson, Jeffrey R.**
Santa Clara, CA 95052-8043 (US)

(74) Representative: **Tollett, Ian et al**
Williams, Powell & Associates,
4 St. Paul's Churchyard Inn
London EC4M 8AY (GB)

(54) Method of producing nucleic acid molecules with reduced secondary structure

(57) The present invention provides a system for generating nucleic acid molecules having reduced levels of secondary structure compared to nucleic acid molecules of the same nucleotide sequence containing only naturally-occurring bases. Such molecules are referred to herein as "unstructured nucleic acids" (UNAs). UNAs have reduced levels of secondary structure because of their reduced ability to form intramolecular hydrogen bond base pairs between regions of substantially complementary sequence. Preferred UNAs, however, retain the ability to form intermolecular hydrogen bond base pairs with other nucleic acid molecules.

Figure 1



EP 1 072 679 A3

Printed by Jouve, 75001 PARIS (FR)

European Patent Office

EUROPEAN SEARCH REPORT

Application Number EP 00 30 0[illegible]

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
A	US 5 432 272 A (BENNER STEVEN A) 11 July 1995 (1995-07-11) * the whole document *		C12N15/11 C12P19/34 C12Q1/68 C07H21/00 //C12N9/12
A	CHOLLET A ET AL: "DNA CONTAINING THE BASE ANALOGEU 2-AMINOADENINE: PREPARATION, USE AS HYBRIDIZATION PROBES AND CLEAVAGE BY RESTRICTION ENDONUCLEASES" NUCLEIC ACIDS RESEARCH, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 16, no. 1, 11 January 1988 (1988-01-11), pages 305-317, XP002024258 ISSN: 0305-1048 * the whole document *		
A	LEBEDEV Y ET AL: "Oligonucleotides containing 2-aminoadenine and 5-methylcytosine are more effective as primers for PCR amplification than their nonmodified counterparts" GENETIC ANALYSIS: BIOMOLECULAR ENGINEERING, ELSEVIER SCIENCE PUBLISHING, US, vol. 13, no. 1, 1 May 1996 (1996-05-01), pages 15-21, XP004052069 ISSN: 1050-3862 * the whole document *		TECHNICAL FIELDS SEARCHED (Int.Cl.7) C12N C12P C12Q C07H
A	BLOOM LINDA B ET AL: "Influence of 5'-nearest neighbors on the insertion kinetics of the fluorescent nucleotide analog 2-aminopurine by Klenow fragment." BIOCHEMISTRY, vol. 32, no. 41, 1993, pages 11247-11258, XP002200596 ISSN: 0006-2960 * the whole document *		
	-/--		

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	31 May 2002	Hornig, H

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 00 30 6128

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
A	FREY MICHELLE WEST ET AL: "The nucleotide analog 2-aminopurine as a spectroscopic probe of nucleotide incorporation by the Klenow fragment of Escherichia coli polymerase I and bacteriophage T4 DNA polymerase." BIOCHEMISTRY, vol. 34, no. 28, 1995, pages 9185-9192, XP002200597 ISSN: 0006-2960 * the whole document *		
A	BAZIN H ET AL: "Reinvestigation of 4-thiothymidine-5'-triphosphate synthesis." NUCLEOSIDES & NUCLEOTIDES, vol. 18, no. 4-5, April 1999 (1999-04), pages 965-966, XP002200598 ISSN: 0732-8311 * the whole document *		TECHNICAL FIELDS SEARCHED (Int.Cl.7)
A	HOFER B ET AL: "Enzymatic synthesis, ligation, and restriction of DNA containing deoxy-4-thiothymidine." NUCLEIC ACIDS RESEARCH. ENGLAND 25 FEB 1981, vol. 9, no. 4, 25 February 1981 (1981-02-25), pages 753-767, XP002200599 ISSN: 0305-1048 * the whole document *		
	-/--		

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	31 May 2002	Hornig, H

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

European Patent Office

EUROPEAN SEARCH REPORT

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim
P,A	RAO T V S ET AL: "Incorporation of 4-thiothymidine into DNA by the klenow fragment and HIV-1 reverse transcriptase" BIOORGANIC & MEDICINAL CHEMISTRY LETTERS, OXFORD, GB, vol. 10, no. 9, May 2000 (2000-05), pages 907-910, XP004199040 ISSN: 0960-894X * the whole document *	
A	TRINH T Q ET AL: "PREFERENTIAL DNA SECONDARY STRUCTURE MUTAGENESIS IN THE LAGGING STRAND OF REPLICATION IN E.COLI" NATURE, MACMILLAN JOURNALS LTD. LONDON, GB, vol. 352, 1991, pages 544-547, XP002926095 ISSN: 0028-0836 * the whole document *	
A	AMBARTSUMY N S ET AL: "ELIMINATION OF THE SECONDARY STRUCTURE EFFECT IN GEL SEQUENCING OF NUCLEIC ACIDS" FEBS LETTERS, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL, vol. 114, no. 2, June 1980 (1980-06), pages 265-268, XP002905760 ISSN: 0014-5793 * the whole document *	

TECHNICAL FIELDS SEARCHED (Int.Cl.7)

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	31 May 2002	Hornig, H

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

CLAIMS:

1. A method of producing an unstructured nucleic acid, comprising the steps of:

i) providing a nucleic acid template including a first nucleic acid sequence and a second nucleic acid sequence substantially complementary to said first nucleic acid sequence;

ii) providing nucleic acid precursors to produce said unstructured nucleic acid which is complementary with said first and second nucleic acid sequences, said nucleotide precursors being unable to hybridise with complementary nucleotide precursors; and

iii) contacting said nucleic acid template and said nucleotide precursors with an enzyme under conditions such that said unstructured nucleic acid is produced.

2. The method according to claim 1, wherein the nucleic acid precursors comprise at least two nucleotides capable of being incorporated enzymatically into a polynucleotide, and wherein said at least two nucleotides are unable to form intramolecular base pairs but can form intermolecular base pairs.

3. The method of claim 1 or claim 2, wherein the sequences complementary to said first nucleic acid sequence and said

one another.

4. The method according to any one of the preceding claims, wherein step ii) comprises providing:

a) 2-aminodeoxyadenosine 5'-triphosphate, and 2-thiodeoxythymidine 5'-triphosphate; or

b) 2-aminodeoxyadenosine 5'-triphosphate, 2-thiodeoxythymidine 5'-triphosphate, deoxyguanosine 5'-triphosphate, and deoxycytidine 5'-triphosphate; or

c) inosine 5'-triphosphate, and pyrrolopyrimidine 5'-triphosphate;

d) 2-aminodeoxyadenosine 5'-triphosphate, 2-thiodeoxythymidine 5' -triphosphate, deoxyinosine 5'-triphosphate, and deoxypyrrolopyrimidine 5'-triphosphate; or

e) deoxyguanosine 5'-triphosphate, and 2-thiodeoxycytidine; or

f) deoxyadenosine 5'-thiodeoxycytidine 5'-triphosphate; or

g) 2-aminoadenosine, 2-thiothymidine, inosine, and pyrrolopyrimidine; or

h) 2-aminodeoxyadenosine 5'-triphosphate, 2-thiodeoxythymidine 5'-triphosphate, deoxyguanosine 5'-triphosphate, and 2-thiodeoxycytidine 5'-triphosphate; or

i) mixtures thereof.

5. The method according to any one of the preceding

claims, wherein said enzyme is selected from the group consisting of: RNA polymerase, DNA polymerase, reverse transcriptase, ribozymes, self-replicating RNA molecules and mixtures thereof.

6. An unstructured nucleic acid with reduced secondary structure relative to a nucleic acid of substantially identical nucleotide sequence having naturally occurring bases, wherein the unstructure nucleic acid has at least two sequence elements that are complementary, wherein the complementary sequence elements do not form intramolecular base pairs, wherein at least one sequence element of at least two complementary sequence elements is capable of hybridizing to a substantially complementary sequence in another nucleic acid.

7. The unstructured nucleic acid of claim 6, wherein the nucleic acid is synthesized by an enzyme, and optionally the enzyme is selected from the group consisting of: RNA polymerase, DNA polymerase, reverse transcriptase, ribozymes and self-replicating RNA molecules.

8. The unstructured nucleic acid of claim 6 or claim 7, wherein the nucleic acid is at least 110 nucleotides in length.

9. The unstructured nucleic acid of claim 6, claim 7 or

claim 8, wherein nucleic acid comprises 2-aminoadenine and 2-thiothymidine.

10. The unstructured nucleic acid of any one of claims 6 to 9 producible by the method of any one of claims 1 to 5.

US005432272A

United States Patent [19]

Benner

[11] **Patent Number: 5,432,272**

[45] **Date of Patent: Jul. 11, 1995**

[54] **METHOD FOR INCORPORATING INTO A DNA OR RNA OLIGONUCLEOTIDE USING NUCLEOTIDES BEARING HETEROCYCLIC BASES**

[76] Inventor: **Steven A. Benner**, Hadlaubstrasse 151, CH-8006 Zurich, Switzerland

[21] Appl. No.: **594,290**

[22] Filed: **Oct. 9, 1990**

[51] Int. Cl.[6] C07H 23/00; C12P 19/34

[52] U.S. Cl. 536/25.3; 536/25.33; 435/91.1; 435/91.51; 435/9.41

[58] Field of Search 536/25, 23, 24, 26, 536/27, 29, 28, 25.3, 25.33; 435/91, 91.1, 91.51, 91.41

[56] References Cited

U.S. PATENT DOCUMENTS

4,959,463	9/1990	Froehler et al.	536/27
5,047,519	9/1991	Hobbs, Jr. et al.	536/23

OTHER PUBLICATIONS

Langer et al. (1981) PNAS vol. 78, pp. 6633–6637.
Cech et al. (1986) Ann. Rev. Biochem. vol. 55, p. 599.
Szostak (1986) Nature 332, 83.
Been et al. (1988) Science 239, 1412.
Leary et al. (1983) PNAS 80, 4045.
Chen et al. (1988) Biochem 27, 6032.
Haseloff et al. (1988) Nature 334, 585.
Sarver et al. (1990) Science 247, 1222–1225.
Noren et al. (1989) Science 244, 182.
Bain et al. (1989) J. Am. Chem. Soc. 111, 8013–8014.
Rich (1962) Horizons in Biochemistry, Kasha et al. editors, N.Y., Academic Press, 103–126.
Saenger (1985) Nucleic Acid Chemistry, Springer—Verlag.
Zubay (1988) The Roots of Modern Biochemistry, Kleinkauf et al. Berlin, Walter de Grumpton & Co., 911–916.
Switzer et al. (1989) J. Am. Chem. Soc. 111, 8322.
Piscisilli et al. (1990) Nature 343, 33–37.
Uhlenbeck (1987), Nature 328, 18.
Chu et al. (1977) J. Org. Chem. 42, 711.
Ugarkar et al. (1984) Heterocyl. Chem. 21, 1865–1870.
Milligan et al. (1987) Nucleic Acids Res. 15, 8783.
Coleman et al. (1988) Biochemistry, 27, 3966.
Roy et al. (1983), Nucleosides and Nuclestides, 2, 237.
Cobianchi et al. (1987) Meth. Enzymol. 152, 94.
Bartholomew et al. (1975) J. Org. Chem. 40, 3708–3713.
Kimura et al. (1980) Bull. Soc. Chem. Jap. 53, 3670.
Mantsch et al. (1975) Biochemistry 14, 5593.
Sgaramella et al. (1972), J. Mol. Biol., 72, 427.
Sanger et al. (1975), J. Mol. Biol., 94, 441.

Primary Examiner—Stephen G. Walsh
Assistant Examiner—Gian P. Wang

[57] **ABSTRACT**

The disclosure describes a method for incorporating into double stranded DNA and RNA base pairs composed of pairing units that fit the Watson-Crick geometry in that they involve a monocyclic six membered ring pairing with a fused bicyclic heterocyclic ring system composed of a five member ring fused with a six membered ring, with the orientation of the heterocycles with respect to each other and with respect to the backbone chain analogous to that found in DNA and RNA, but with a pattern of hydrogen bonds holding the base pair together different from that found in the AT and GC base pairs (a "non-standard base pair").

7 Claims, 5 Drawing Sheets

Generalized structures for 4 non-standard base pairs formed between 8 non-standard bases

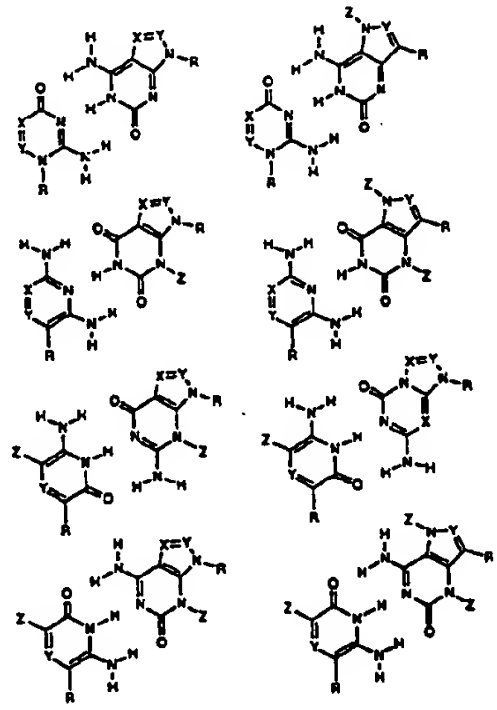


Figure 1: The two natural base pairs formed between the four natural bases

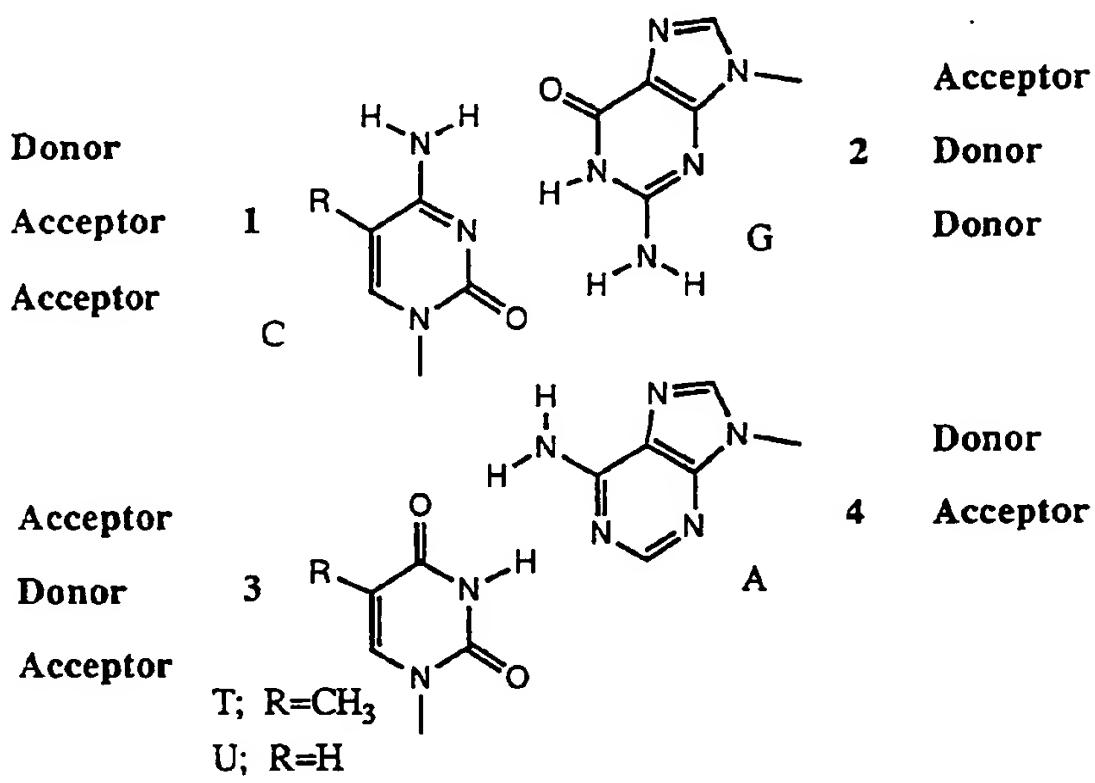


Figure 2: Examples of 4 non-standard base pairs formed between 8 non-standard bases

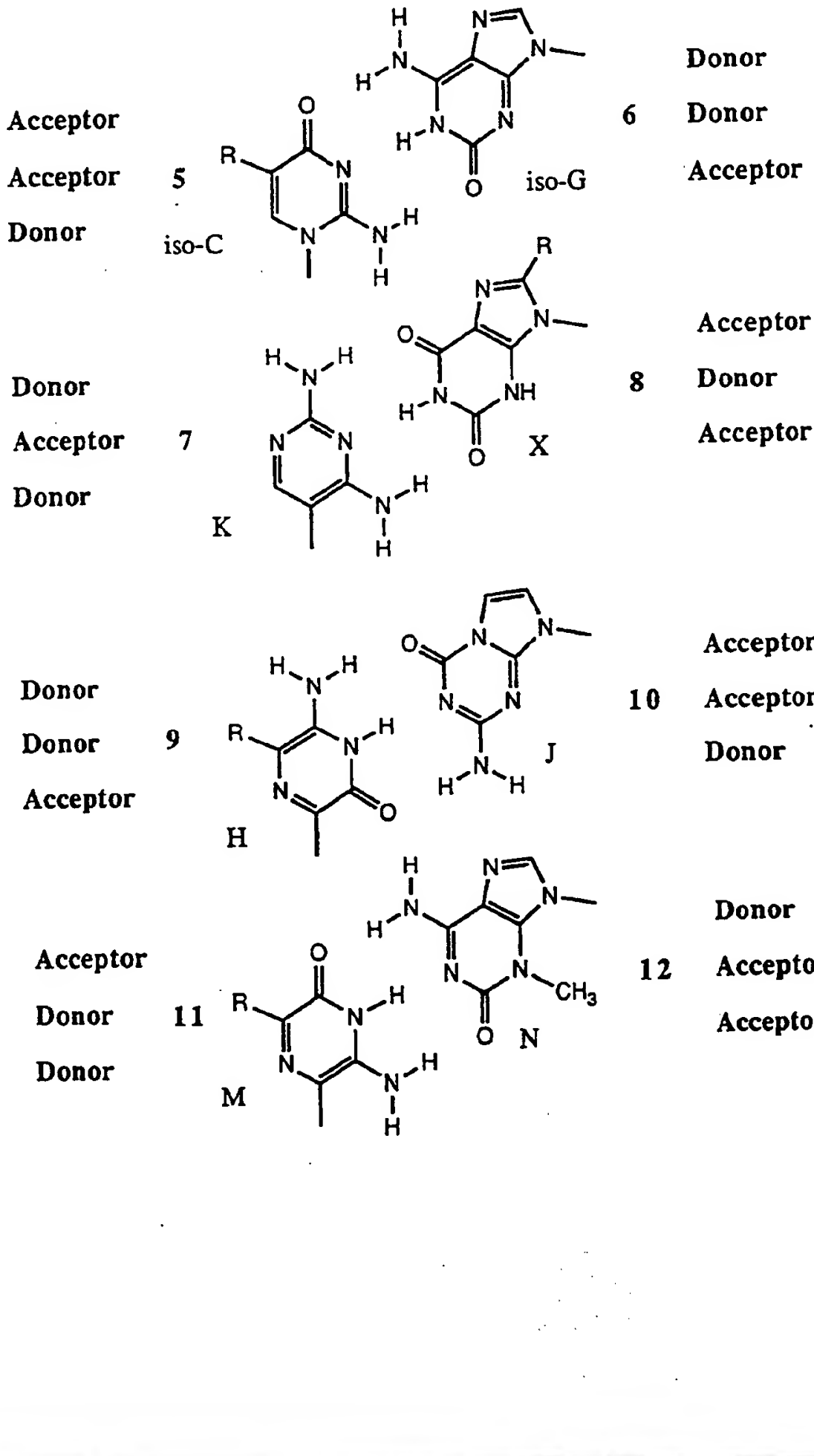


Figure 3: Novel base pair discussed by Zubay

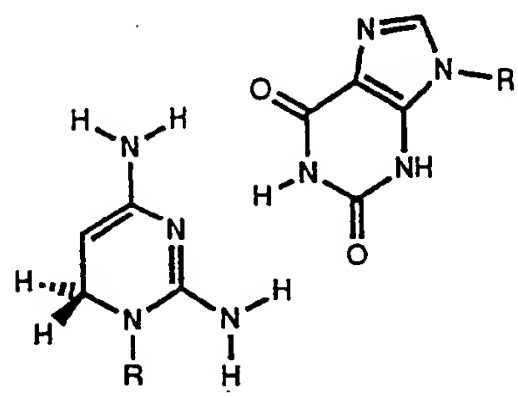


Figure 4: Generalized structures for 4 non-standard base pairs formed between 8 non-standard bases

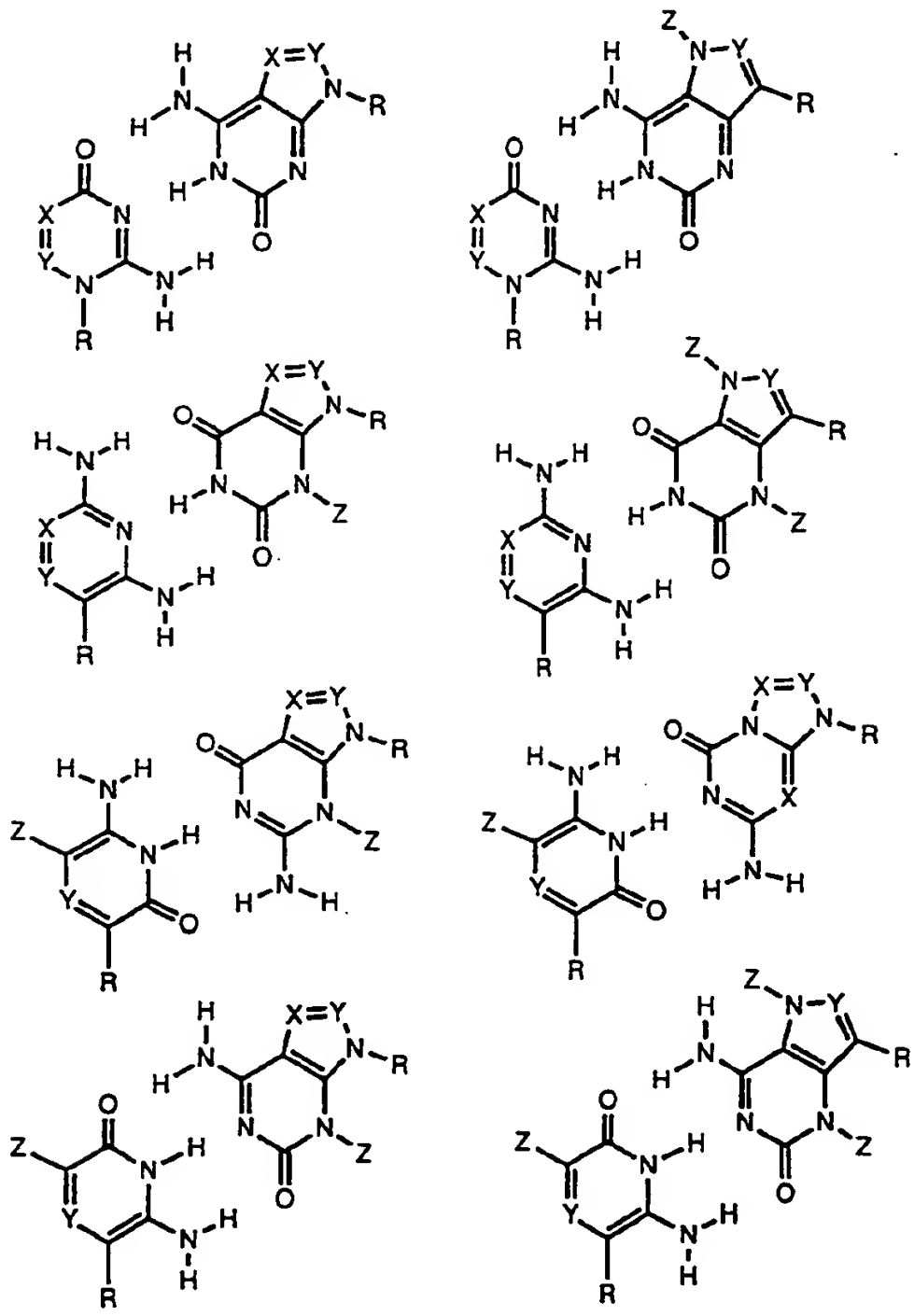


Figure 5: Templates used in the Examples

Templates for T7 RNA polymerase

d-5'-GATTTTGA
d-3'-CTAAAACTGGKGA

d-5'-GATTTTGA
d-3'-CTAAAACTGG*iso*-CGA

d-5'-GATTTTGA
d-3'-CTAAAACTGGTGA

d-5'-GATTTTGA
d-3'-CTAAAACTGGCGA

Templates for DNA polymerase

d-5'-TAATACGACTCACTATAG
d-3'-ATTATGCTGAGTGATATCGCGGCKCGA

d-5'-TAATACGACTCACTATAG
d-3'-ATTATGCTGAGTGATATCGCGGC*iso*-CCGA

d-5'-TAATACGACTCACTATAG
d-3'-ATTATGCTGAGTGATATCGCGGCCCGA

METHOD FOR INCORPORATING INTO A DNA OR RNA OLIGONUCLEOTIDE USING NUCLEOTIDES BEARING HETEROCYCLIC BASES

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

This invention pertains to increasing the number of building blocks that can be incorporated independently into oligonucleotides via DNA and RNA polymerases, and nucleoside analogs capable of forming non-standard Watson-Crick base pairs joined by patterns of hydrogen bonding different from those found in the adenine-thymine and cytosine-guanine base pairs.

SUMMARY OF THE INVENTION

The objective of this invention is to increase the number of independently replicatable building blocks that can be incorporated into DNA and RNA via template directed polymerization. The objective is accomplished by a method for incorporating into double stranded DNA and RNA base pairs composed of pairing units that fit the Watson-Crick geometry in that they involve a monocyclic six membered ring pairing with a fused bicyclic heterocyclic ring system composed of a five member ring fused with a six membered ring, with the orientation of the heterocycles with respect to each other and with respect to the backbone chain analogous to that found in DNA and RNA, but with a pattern of hydrogen bonds holding the base pair together different from that found in the AT and GC base pairs (a "non-standard base pair").

BRIEF DESCRIPTION OF FIGURES

FIG. 1: The two natural base pairs formed between the four natural bases

FIG. 2: General structure of four novel base pairs formed between eight novel bases disclosed here

FIG. 3: Novel base pair discussed by Zubay

FIG. 4: Generalized structures

FIG. 5: Templates used in the examples

DESCRIPTION OF THE PREFERRED EMBODIMENTS

Natural oligonucleotides bind to complementary oligonucleotides according to the well-known rules of base pairing first elaborated by Watson and Crick, where adenine (A) pairs with thymine (T) or uracil (U), and guanine (G) pairs with cytosine (C), with the complementary strands anti-parallel to one another. These pairing rules allow for the specific hybridization of an oligonucleotide with complementary oligonucleotides, making oligonucleotides valuable as probes in the laboratory, in diagnostic applications, as messages that can direct the synthesis of specific proteins, and in a wide range of other applications well known in the art. Further, the pairing is the basis by which enzymes are able to catalyze the synthesis of new oligonucleotides that are complementary to template nucleotides. In this synthesis, building blocks (normally the triphosphates of ribo or deoxyribo derivatives of A, T, U, C, or G) are directed by a template oligonucleotide to form a complementary oligonucleotide with the correct sequence. This process is the bases for replication of all forms of life, and also serves as the basis for all technologies for enzymatic synthesis and amplification of specific heterosequence nucleic acids by enzymes such as DNA and RNA polymerase, and in the polymerase chain reaction.

The Watson-Crick pairing rules can be understood chemically in terms of the arrangement of hydrogen bonding groups on the heterocyclic bases of the oligonucleotide, groups that can either be hydrogen bond donors or acceptors (FIG. 1). In the standard Watson-Crick geometry, a large purine base pairs with a small pyrimidine base; thus, the AT base pair is the same size as a GC base pair. This means that the rungs of the DNA ladder, formed from either AT or GC base pairs, all have the same length.

Further recognition between bases is determined by hydrogen bonds between the bases. Hydrogen bond donors are heteroatoms (nitrogen or oxygen in the natural bases) bearing a hydrogen; hydrogen bond acceptors are heteroatoms (nitrogen or oxygen in the natural bases) with a lone pair of electrons. In the geometry of the Watson-Crick base pair, a six membered ring (in natural oligonucleotides, a pyrimidine) is juxtaposed to a ring system composed of a fused six membered ring and a five membered ring (in natural oligonucleotides, a purine), with a middle hydrogen bond linking two ring atoms, and hydrogen bonds on either side joining functional groups appended to each of the rings, with donor groups paired with acceptor groups (FIG. 1).

Derivatized oligonucleotide building blocks, where a side chain has been appended to one of the nucleoside bases A, T, U, G, or C (the "normal" bases), have application because of their combination of Watson-Crick base pairing and special reactivity associated with the chemical properties of the side chain. For example, oligonucleotides containing a T to which is appended a side chain bearing a biotin residue can first bind to a complementary oligonucleotide, and the hybrid can then be isolated by virtue of the specific affinity of biotin to avidin (Langer, P. R.; Waldrop, A. A.; Ward, D. C. (1981) *Proc. Nat. Acad. Sci.* 78, 6633–6637), and finds application in diagnostic work. Oligonucleotides containing special functional groups (e.g., thiols or hydrazines) can be immobilized to solid supports more readily than those composed solely of the five "natural" bases.

Often, derivatized building blocks can be incorporated into oligonucleotides by enzymatic transcription of natural oligonucleotide templates in the presence of the triphosphate of the derivatized nucleoside, the substrate of the appropriate (DNA or RNA) polymerase. In this process, a natural nucleoside is placed in the template, and standard Watson-Crick base pairing is exploited to direct the incoming modified nucleoside opposite to it in the growing oligonucleotide chain.

However, the presently available base pairs are limited in that there are only two mutually exclusive hydrogen bonding patterns available in natural DNA. This means that should one wish to introduce a modified nucleoside based on one of the natural nucleosides into an oligonucleotide, it would be incorporated wherever the complementary natural nucleoside is found in the template. For many applications, this is undesirable.

Many of the limitations that arise from the existence of only four natural nucleoside bases, joined in only two types of base pairs via only two types of hydrogen bonding schemes, could be overcome were additional bases available that could be incorporated into oligonucleotides, where the additional bases presented patterns of hydrogen bond donating and accepting groups in a pattern different from those presented by the natural

bases, and therefore could form base pairs exclusively with additional complementary bases. The purpose of this invention is to describe compositions of matter containing these additional bases, and methods for their incorporation into analogs of oligonucleotides.

In the naturally-occurring base pairs, the pyrimidines components present an acceptor-donor-acceptor (T) or a donor-acceptor-acceptor (C) pattern of hydrogen bonds to a purine on an opposite strand. This invention is based on the fact that bases with other patterns of hydrogen bond donating and accepting groups can fit the standard Watson-Crick geometry. For example, FIG. 2 discloses four base pairs that have still different patterns, an acceptor-acceptor-donor pattern for iso-C, and donor-acceptor-donor pattern for K. Bases, pairing schemes, and base pairs that have hydrogen bonding patterns different from those found in the AT and GC base pairs are here termed "non-standard". Although not found (so far) in Nature, the non-standard base pairs shown in FIG. 2) apparently can fit into the DNA ladder in a standard Watson-Crick duplex.

Further, the patterns of hydrogen bonds in these non-standard pyrimidines are different from each other, and different from those in the natural pyrimidines T and C. This suggested that in an enzyme-catalyzed polymerization, it might be possible for each non-standard pyrimidine to recognize uniquely its complementary purine with high fidelity. Thus, it should be possible to make copies of a DNA molecule containing all 12 bases simply by following an expanded set of Watson-Crick rules: A pairs with T, G pairs with C, iso-C pairs with iso-G, and K pairs with X, H pairs with J, and M pairs with N (FIG. 2). In other words, it should be possible to have a genetic alphabet with twelve bases instead of four.

Statements considering non-standard base pairs in a general way can, to our knowledge, be found only three times previously in the literature. Considering possible bases that might have been incorporated into nucleic acids in the first forms of life on the earth two to four billion years ago, Rich mentions the base pair between iso-G and iso-G (Rich, A. (1962), Horizons in Biochemistry, Kasha, M. and Pullman, B. editors, N.Y., Academic Press, 103–126) as a base pair that was conceivable, but rejected, by the earliest forms of life. However, Rich did not disclose nor make obvious the process disclosed here employing contemporary DNA and RNA polymerases as part of a process for incorporating the base pair between iso-G and iso-G into oligonucleotides. Saenger (Saenger, W. (1988) *Principles of Nucleic Acid Structure* New York, Springer-Verlag, pages 114–115) also mentions this base pair, but concludes, based on the fact that iso-G has a alternate tautomeric forms (vide infra), that it has no utility as part of an oligonucleotide that is to be copied.

Zubay (Zubay, G. (1988) *The Roots of Modern Biochemistry*, Kleinkauf, von Doehren, Jaenicke, Berlin, Walter de Gruyter & Co. 911-916) suggested that 2,4-diamino-5,6-dihydropyryimidine-1-riboside, with a donor-acceptor-donor pattern, might be able to pair with xanthosine (FIG. 3). In Zubay's suggested pyrimidine, however, the pyrimidine ring is not aromatic and therefore not planar. Although it has never been examined experimentally, we believe on these grounds that it would not participate well in "base stacking," the interaction (vide supra) that is important for the stability of a double helix. Further, Zubay's base incorporates the structural unit known as a "vinylogous enamine", a structural unit that is likely to be unstable in acidic solution. Thus, we doubt that it can be incorporated into an oligonucleotide by enzymatic transcription of a complementary oligonucleotide.

Zubay discloses neither experimental studies with his suggested base nor the potential utility of a new base pair that would arise were the new base a substrate for DNA and RNA polymerases present in the modern world. Further, the possibility of constructing additional base pairing schemes (such as the non-standard base pairs disclosed in FIG. 2), was explicitly denied. Zubay writes "We have searched for other purine-pyrimidine base pairs with different arrangements of hydrogen bonding groups that would satisfy the criterion of exclusive pairing. No additional pairs have been found. Thus except for modifications at non-hydrogen bonding sites the additional base pair described here may be unique. "This comment from a prominent figure in American biochemistry supports the notion that the invention disclosed here, where DNA and RNA polymerases can be used in a method to incorporate non-standard base pairs into oligonucleotides, is not obvious to one skilled in the art.

Should the additional base pairs disclosed in FIG. 2 be incorporated enzymatically into DNA and RNA, they could be useful for a variety of purposes. For example, RNA molecules prepared by transcription, although it is known to be a catalyst under special circumstances (a) Cech, T. R.; Bass, B. L. *Ann. Rev. Biochem.* 1986, 55, 599. (b) Szostak, J. W. *Nature* 1986, 332, 83. (c) Been, M. D.; Cech, T. R. *Science* 1988, 239, 1412), appear to have a much smaller catalytic potential then proteins because they lack building blocks bearing functional groups. Conversely, the limited functionality present on natural oligonucleotides constrains the chemist attempting to design catalytically active RNA molecules, in particular, RNA molecules that catalyze the template-directed polymerization of RNA.

Additional base pairs could relax these constraints, especially if their hydrogen bonding pattern differed from those in the AT and GC base pairs, as novel hydrogen bonding schemes would allow additional base pairs to be incorporated enzymatically at specific positions in an oligonucleotide molecule (Switzer, C. Y, Moroney, S. E. & Benner, S. A. *J. Am. Chem. Soc.*, 1989, 111, 8322). If functionalized, such additional bases should also allow the incorporation by transcription of functional groups directly into RNA; the remaining unfunctionalized building blocks could then control secondary structure without introducing over-functionalization and attendant non-specific catalysis. Further, bases bearing functional groups at the position structurally analogous to the 5-position of the uridine ring should be substrates for most polymerases (Leary, J. L., Brigati, D. J. & Ward, D. C. *Proc. Natl. Acad. Sci.* 1983, 80, 4045). New base pairs should also find use in studies of the structure of biologically important RNA and DNA molecules (Chem, T. R., Churchill, M. E. A. Tullius, T. D. Kallenbach, N. R., Seemann, N. C. (1988) *Biochem.*, 27, 6032) and protein-nucleic acid interactions. Several types of catalytic RNA molecules containing natural bases have been proposed as anti-viral agents, for use in agriculture, and in other areas. (Haseloff, J., Gerlach, W. L. *Nature*, 1988, 334, 585; Sarver, N., Cantin, E. M., Chang, P. S., Zaia, J. A., Ladne, P. A., Stephens, D. A., Rossi, J. J. *Science*, 1990, 247 1222–1225). Catalytic RNA molecules incorporating additional bases should be even more useful in certain of

these applications. A segment of DNA or RNA containing the additional bases could be replicated only in the presence of triphosphates of the complementary additional bases, allowing the selective copying of DNA containing the additional bases in the presence of DNA containing normal bases, and vice versa. More speculatively, the extra letters in the nucleoside alphabet might eventually be used to expand the genetic code, increasing the number of amino acids that can be incorporated translationally into proteins Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C. & Schultz, P. G. *Science*, 1989, 244, 182; J. D. Bain, J. C. G. Glabe, T. A. Dix, A. R. Chamberlain *J. Am. Chem. Soc.* 1989, 111, 8013–8014.

The process disclosed here involves the incubation of an oligonucleotide template containing one or more non-standard bases with either a DNA or RNA polymerase (Uhlenback, O. C. *Nature*, 1987, 328, 18) in the presence of triphosphates of the complementary nucleosides.

Experiments with iso-C and iso-G (FIG. 2) have shown that bases with novel hydrogen bonding schemes can be incorporated by DNA and RNA polymerases. Thus, this base pair fulfills the criteria outlined above. However, certain forms of iso-C were found to hydrolyze slowly to U under conditions of DNA synthesis, introducing into a template a base that pairs with A in the place of a base that pairs with iso-G. Further, iso-G exists to some extent in a minor tautomeric form that is complementary to U and T. Although the existence of a minor tautomeric form of iso-G has some advantages in certain circumstances (e.g., when one wishes to introduce iso-G into an oligonucleotide duplex opposite U or T), these problems complicate the selective incorporation of the iso-C/iso-G base pair in oligonucleotides also containing A and T. Therefore, the search for a preferred embodiment was directed towards base pairs in FIG. 2 where the non-standard base is joined to the sugar by a carbon-carbon bond, where chemical considerations suggested that hydrolysis and tautomeric equilibria might be less problematic.

Much work was directed towards the pyridine nucleoside analog, 3-β-D-ribofuranosyl-(2,6-diaminopyridine). However, the presently preferred base pairs are those where the six-ring pyrimidine analog is joined to a ribose or deoxyribose ring via a carbon-carbon bond, and where the pyrimidine analog contains at least two nitrogens in the ring itself. Give an appropriate arrangements of hydrogen bond donating and accepting groups, many ring systems are appropriate (FIG. 4), including ring systems to which are appended functionalized and unfunctionalized side chains. Syntheses for many of these compounds are known in the prior art. (Bartolomew, D. G., Dea, P., Robins, R. K. Revenkar, G. R. (1975) *Org. Chem.*, 40, 3708). However, there are several constraints on the ring system and its substituents. First, the ring systems must be aromatic so that they are capable of stacking with base above and below in the double helix. Second, substituent on the 6 position of the pyrimidine (or the analogous position of a pyrimidine analog) and the 8 position of the purine (or the analogous position of a purine analog) is preferably no larger than hydrogen.

The most preferred pyrimidine analog is 3-β-D-ribofuranosyl-(2,6-diaminopyrimidine), trivially designated here as K. Several complementary purines are presently preferred. For example, either xanthosine or 3β-D-ribofuranosyl-(1-methyl-pyrazolo[4,3-d]pyrimidine-5,7(4H,6H)-dione), also known as 7-methyl oxoformycin B, have been found to form base pairs with K when incorporated into DNA and RNA.

EXAMPLE 1

The K-P Base Pair

The pyrimidine 3β-D-ribofuranosyl-(2,6-diaminopyrimidine), trivially designated as K, presents a donor-acceptor-donor hydrogen bonding pattern to a complementary strand in a duplex structure. K as a deoxyriboside derivative suitable for automated DNA synthesis was synthesized from a known precursor by routes known in the prior art. C. K. Chu, U. Reichman, K. A. Watanable, J. J. Fox, *J. Org. Chem.* 1977, 42, 711. Two purine analogs were chosen to complement K. The first, xanthosine (X), is a natural base available commercially as both the nucleoside and nucleoside triphosphate. However, because of concerns that deoxyxanthosine might undergo depurination in some of the studies planned, another complementary base, 3-β-D-ribofuranosyl-(1-methyl-pyrazolo[4,3-d]pyrimidine-5,7(4H,6H)-dione) (B. G. Ugarkar, G. R. Revankar, R. K. Robins, *J. Heterocycl. Chem.*, 21, 1965–1870 (1984)), also known as 7-methyl oxoformycin B, and trivially designated here as P, was prepared by routes known in the prior art. In P, the heterocyclic base is joined to the pentose ring by a carbon-carbon bond.

The K-P base pair has physical and chemical properties suitable as replicatable components of a genetic alphabet (J. A. Piccirilli, T. Krauch, S. E. Moroney, S. A. Benner, *Nature*, 343, 33–37 (1990). In a solution of a derivative of K in chloroform, a strong nuclear Overhauser enhancement (NOE) between the proton at C1' of the ribose ring and the proton at C4 of the heterocyclic ring suggested that K adopts the undesired syn conformation when alone in solution. However upon addition of a protected derivative of the complementary purine nucleoside P, this NOE largely disappeared. Further, when both K and its complement are present, the resonances assigned to the amine protons of K shift strongly downfield, as does the resonance of P assigned to the nitrogen flanked by the carbonyl groups. These facts together show that K and P form a standard Watson-Crick base pair in a solution of chloroform.

To determine the effect of a base pair between K and P on the stability of a DNA duplex, several oligonucleotides containing the K-P base pair were synthesized using an Applied Biosystems Oligonucleotide Synthesizer. Melting studies showed that duplexes containing a K-P base pair are only slightly less stable than duplexes containing only natural bases. Further, duplexes containing the new base pair are considerably more stable than those containing mismatches involving the new bases, which in turn had melting temperatures similar to duplexes containing mismatches of natural bases (Piccirilli et al., op. cit.). The stability of various mismatches was consistent with the presumed stability of "wobble" base pairs, which should be particularly important for the GT and AK mismatches. These results suggested that enzymatic incorporation of a new base selectively opposite its complement in a DNA template should be possible, provided that natural DNA and RNA polymerases accepted the new bases.

To demonstrate that xanthosine triphosphate could be incorporated enzymatically into an RNA oligonucleotide opposite a K in the template, a promoter-template

including a promoter sequence of T7 RNA polymerase (Milligan, J. F., Groebe, D. R., Witherell, G. W., & Uhlenbeck, O. C. *Nucleic Acids Res.*, 15, 8783 (1987), (17 bases) followed by a short oligonucleotide segment (7 bases), the new base K, 1 additional base, and a final A was synthesized, together with a complementary 18 base primer (FIG. 5). Incorporation of K into the synthetic DNA templates was verified by digestion of samples of the template with snake venom phosphodiesterase, hydrolysis of the phosphate from the products with bacterial alkaline phosphatase, and analysis of the resulting nucleosides by HPLC (data not shown). Control templates containing T replacing K were also prepared by synthesis. Transcription of the primed templates could be detected most simply by the incorporation of radiolabled UMP (from a-labeled UTP) into a product RNA molecule 10 bases long (the "full length product").

When synthetic template 1 was incubated with labeled UTP and various other nucleoside triphosphates, full length products were observed in the presence of XTP. The efficiency of synthesis of full length product from templates with and without K was approximately the same, provided that the necessary complementary nucleoside triphosphates were all present in the incubation mixtures. In absence of XTP, a significant amount of full length product could be detected only in the presence of ATP, and this at somewhat low levels (ca. 24%, measured by scintillation counting of bands cut from the gel). Such a misincorporation presumably occurs via "wobble" base pairing, and is not infrequent even with natural bases when incorporation experiments are run in incubation mixtures that are missing one component. Coleman, J. E., Martin, C. T. & Muller, D. K. *Biochemistry*, 27, 3966 (1988).

To determine whether misincorporation of A was diminished by competition with X, experiments were performed with tritiated XTP (synthesized from tritiated GTP by Demijanov oxidation) Roy, K. B. & Miles, H. T. *Nucleosides and Nucleotides*, 2, 237 (1983) and γ-^{32}P-labeled GTP (which is incorporated, with the triphosphate intact, at the 5' end of the RNA product) together in an incubation mixture in varying ratios. Full length products from an incubation containing a 1:1 molar ratio of ^{3}H-XTP and unlabeled ATP were isolated by gel electrophoresis, the bands excised, and the relative amounts of ^{3}H and ^{32}P determined by scintillation counting. After correction for the specific activities of the starting nucleotides, the misincorporation of adenosine into the product at a XTP:ATP ratio of 1:1 was reduced to ca. 14%. Infidelity further decreases with increasing ratios of X:A, and most likely stems from errors made by the polymerase itself rather than from minor tautomers of the bases.

To demonstrate that xanthosine triphosphate could be incorporated into a DNA oligonucleotide opposite a K in the template, a set of primer-templates (FIG. 5) were prepared containing either K, C or T (the latter two serving as control templates). Incorporation of K into the synthetic DNA templates was again verified by digestion of samples of the template with snake venom phosphodiesterase, removal of the phosphate from the products by bacterial alkaline phosphatase, and analysis of the resulting nucleosides by HPLC. As before, the last base in the template was a unique A, permitting the detection of full length products most simply by autoradiography following the incorporation of α-^{32}P-TTP.

The synthetic primer-templates were incubated with the Klenow fragment of DNA polymerase I (Pol 1) Cobianchi, F. & Wilson, S. H. *Meth. Enzymol.*, 152, 94 (1987) in the presence of various nucleoside triphosphates, and the products analyzed by gel electrophoresis. K in the template directed the incorporation of XTP into full length product. Upon electrophoresis, the product containing X migrates faster than the analogous products containing G or A, presumably because the xanthine heterocycle carries an additional negative charge under the conditions of the electrophoresis due to its low pK_a (pK_a=5.7). Direct evidence for the incorporation of xanthosine was obtained by digestion of the product oligonucleotide, kinasing, and electrophoretic analysis.

To measure the relative efficiency as templates of the oligonucleotides containing different bases, product bands from electrophoresis gels were excised and their radioactivity determined by liquid scintillation counting. Templates containing K were ca. 70% as efficient at directing the synthesis of full length product (in the presence of XTP) as those containing only natural bases.

The fidelity of incorporation of X opposite K was examined by incubating templates containing C, T and K with purine triphosphates separately and in competition (FIG. 5). As expected, the fidelity of incorporation was considerably higher with DNA polymerase than with T7 RNA polymerase. Essentially no G or A was incorporated by the Klenow fragment of DNA polymerase opposite K, and essentially no X was incorporated opposite T in the template. The only evidence of infidelity was a low level (ca. 5%) of X misincorporated opposite C in the template when GTP was missing from the incubation mixture. This misincorporation was not observed at all when GTP and XTP were present in a 1:1 ratio.

EXAMPLE 2

The isoC-isoG Base Pair

Protected d-iso-C suitable as a building block for the chemical synthesis of DNA was synthesized by direct extensions of standard methods. Watanabe, K. A.; Reichman, C. K.; Fox, J. J. *Nucleic Acid Chemistry;* Tipson, R. S.; Townsend, L. B., Eds.; John Wiley and Sons: New York 1978; Part 1, p 273. (b) Kimura, J.; Yagi, K.; Suzuki, H.; Mitsunobu, O. *Bull. Soc. Chem. Jap.* 1980, 53, 3670. N^2-benzoyl-5'-dimethoxytrityl-d-iso-C diisopropyl phosphoramidite, used directly in machine-DNA synthesis, was synthesisized from d-iso-C by the general procedure of Atkinson and Smith: Atkinson, T.; Smith, M. *Oligonucleotide Synthesis: A Practical Approach;* Gait, M. J. Ed.; IRL Press: Oxford 1985; pp 35–82. This was incorporated into two templates, and three other templates containing only natural bases were synthesized for use as standards and controls. An 8-mer primer was annealed to the appropriate templates (FIG. 5) to provide a double stranded binding site for the Klenow fragment of DNA polymerase I (*E. coli*), followed by a single stranded coding region containing d-iso-C flanked only by purine nucleotides. Alternatively, different templates (FIG. 5) were annealed to an 18-mer to give the double stranded promoter region required by T7 RNA polymerase, followed by a single stranded coding region containing d-iso-C. In all of the templates, a unique A at the end of the coding strand was included to direct the incorpora-

tion of radiolabelled T or U and ribo- and deoxyribo-iso-GTP's. Mantsch, H. H. et. al., *Biochemistry* 1975, 14, 5593. The reactions with the Klenow fragment were conducted by incubating template/primer, polymerase, and a mixture of the required dNTPs including (α-^{32}P)TTP. Following incubation, the products were analyzed by gel electrophoresis and autoradiography. With primed templates containing iso-C, full length product was obtained only with d-iso-GTP present in the incubation mixture. The presence of iso-G at the correct position in the product oligonucleotide was positively established by a "nearest neighbor" analysis, Sgaramella, V.; Khorana, H. G. *J. Mol. Biol.* 1972, 72, 427. and by the "minus" sequencing method. Sanger, F.; Coulson, A. R. *J. Mol. Biol.* 1975, 94, 441. As expected, in an incubation of a primed template containing T with dATP and the required dNTPs in the absence of d-iso-GTP, full length product was observed only to the extent anticipated by the fact that a small amount (15%) of dUTP was present in the template due to the deamination of iso-C (vide supra).

Infidelity between iso-G and T was anticipated due to the known existence of a minor "phenolic" tautomer of iso-G in addition to the major N_1-H tautomer (Sepiol, J., Kazimierczuk, Z., Shugar, D. Z. *Naturforsch.*, 1976, 31c, 361; the possibility that this minor tautomer could form a Watson-Crick base pair with T was recognized on theoretical grounds. In fact, incubation of a primed template containing T in place of d-iso-C with the required dNTPs and d-iso-GTP did yield a significant amount of full length product. This result strongly suggests that polymerases synthesize a base pair between T and the "phenolic" tautomer of iso-G. This fact diminishes the value of the base pair between iso-G and iso-C for many (but not all applications.

In analogous experiments, T7 RNA polymerase was shown to accept the new base pair. Incubation of a template (FIG. 5) possessing the T7 promoter with the required NTPs yielded more full length product in the presence iso-GTP than in its absence. Sequencing of the RNA transcript positively established the presence of iso-G in the product at the expected position.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i i i) NUMBER OF SEQUENCES: 6

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE:
(A) DESCRIPTION: nucleic acid

(v i) ORIGINAL SOURCE: synthetic

(x) PUBLICATION INFORMATION:
(A) AUTHORS:
Switzer, C. Y.
Moroney , S. E.
Benner, S. A.
(B) TITLE: Enzymatic Incorporation of a New Base Pair into DNA and RNA
(C) JOURNAL: Journal of the American Chemical Society
(D) VOLUME: 111
(F) PAGES: 8322-8323
(G) DATE: 1989

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGTGGTCAAA ATC 13

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE:
(A) DESCRIPTION: nucleic acid

(v i) ORIGINAL SOURCE: synthetic

(x) PUBLICATION INFORMATION:
(A) AUTHORS:
Switzer, C. Y.
Moroney , S. E.
Benner, S. A.

-continued

(B) TITLE: Enzymatic Incorporation of a New Base Pair into DNA and RNA
(C) JOURNAL: Journal of the American Chemical Society
(D) VOLUME: 111
(F) PAGES: 8322-8323
(G) DATE: 1989

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AGCGGTCAAA ATC 13

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE:
(A) DESCRIPTION: nucleic acid

(v i) ORIGINAL SOURCE: synthetic

(x) PUBLICATION INFORMATION:
(A) AUTHORS:
Switzer, C. Y.
Moroney , S. E.
Benner, S. A.
(B) TITLE: Enzymatic Incorporation of a New Base Pair into DNA and RNA
(C) JOURNAL: Journal of the American Chemical Society
(D) VOLUME: 111
(F) PAGES: 8322-8323
(G) DATE: 1989

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TAATACGACT CACTATAG 18

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE:
(A) DESCRIPTION: nucleic acid

(v i) ORIGINAL SOURCE: synthetic

(x) PUBLICATION INFORMATION:
(A) AUTHORS:
Switzer, C. Y.
Moroney , S. E.
Benner, S. A.
(B) TITLE: Enzymatic Incorporation of a New Base Pair into DNA and RNA
(C) JOURNAL: Journal of the American Chemical Society
(D) VOLUME: 111
(F) PAGES: 8322-8323
(G) DATE: 1989

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TAATACGACT CACTATAG 18

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE:
(A) DESCRIPTION: nucleic acid

-continued

(v i) ORIGINAL SOURCE: synthetic

(x) PUBLICATION INFORMATION:
(A) AUTHORS:
Piccirilli, J. A.
Krauch, T.
Moroney , S. E.
Benner, S. A.
(B) TITLE: Extending the Genetic Alphabet: Enzymatic Incorporation of a New Base Pair into DNA and RNA
(C) JOURNAL: Nature
(D) VOLUME: 343
(F) PAGES: 33-37
(G) DATE: 1990

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TAATACGACT CACTATAG 18

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE:
(A) DESCRIPTION: nucleic acid

(v i) ORIGINAL SOURCE: synthetic

(x) PUBLICATION INFORMATION:
(A) AUTHORS:
Switzer, C. Y.
Moroney , S. E.
Benner, S. A.
(B) TITLE: Enzymatic Incorporation of a New Base Pair into DNA and RNA
(C) JOURNAL: Journal of the American Chemical Society
(D) VOLUME: 111
(F) PAGES: 8322-8323
(G) DATE: 1989

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AGCCCGGCGC TATAGTGAGT CGTATTA 27

What is claimed is:

1. A method for incorporating into a DNA or RNA oligonucleotide chain at least one nucleotide unit bearing a heterocyclic base selected from the group consisting of the structural formulae of FIG. 4, wherein —R designates the point of attachment of the base to position 1 of a ribose or deoxyribose ring, X is either a nitrogen atom or a carbon atom bearing a substituent Z, Z is either a hydrogen, an unfunctionalized lower alkyl chain, or a lower alkyl chain bearing an amino, carboxyl, hydroxyl, thiol, aryl, indole, or imidazolyl group, Y is either N or CH, and each ring contains no more than three nitrogens consecutively bonded, that consists of synthesizing a template that is an oligonucleotide containing one or more nucleotide subunits bearing a heterocyclic base selected from the group consisting of the structural formulae of FIG. 4, dissolving this template in buffered aqueous solution, adding to the solution nucleoside triphosphates complementary to the nucleotide subunits in the template, adding to this solution a solution of a DNA or RNA polymerase, and incubating said mixture of the solutions such that an oligonucleotide chain complementary to the template and incorporating said at least one oligonucleotide unit bearing a heterocyclic base selected from the group consisting of the structural formulae of FIG. 4 is synthesized.

2. The method of claim 1 wherein said heterocyclic base is selected from the group consisting of 2,6-diaminopyrimidine, xanthine, 2,6-diaminopyrazine, 1-methyl-pyrazolo[4,3-d]pyrimidine-5,7(4H,6H)-dione, iso-guanine, and iso-cytosine.

3. The method of claim 1 wherein said polymerase is selected from the group consisting of the Klenow fragment of DNA polymerase I, T7 RNA polymerase, and AMV reverse transcriptase.

4. The method of claim 1 wherein said heterocyclic base is iso-cytosine incorporated opposite iso-guanine in a template.

5. The method of claim 1 wherein said heterocyclic base is iso-guanine incorporated opposite iso-cytosine in a template.

6. The method of claim 1 wherein said heterocyclic base is 2,6-diaminopyrimidine or 2,6-diaminopyrazine incorporated opposite 1-methyl-pyrazolo[4,3-d]pyrimidine-5,7(4H,6H)-dione or xanthine in a template.

7. The method of claim 1 wherein said heterocyclic base is 1-methyl-pyrazolo[4,3-d]pyrimidine-5,7(4H,6H)-dione or xanthine incorporated opposite either 2,6-diaminopyrimidine or 2,6-diaminopyrazine in a template.

* * * * *

Volume 16 Number 1 1988 Nucleic Acids Research

DNA containing the base analogue 2-aminoadenine: preparation, use as hybridization probes and cleavage by restriction endonucleases

Andre Chollet*+ and Eric Kawashima+

Department of Chemistry, Biogen SA, Geneva, Switzerland

Received October 15, 1987; Revised November 23, 1987; Accepted December 9, 1987

ABSTRACT

The base analogue 2-aminoadenine (2,6-diaminopurine,D) has been introduced at selected positions into synthetic oligodeoxyribonucleotides and. DNA by the combined use of chemical and enzymatic methods. 2-aminoadenine substitution for adenine introduces changes in the minor groove of DNA and creates an additional hydrogen bond in the Watson-Crick base pair with thymine. Oligonucleotide hybridization probes containing 2-aminoadenine showed increased selectivity and hybridization strength during DNA-DNA hybridization to phage or genomic target DNA. Properties of the base analogue with respect to DNA modifying enzymes were examined. 2-aminoadenine was used to probe minor groove determinants during the treatment of DNA by 12 restriction endonucleases. Inhibition of cleavage was found for several restriction enzymes.

INTRODUCTION

The substitution of bases by analogues in DNA is often a useful approach to investigate DNA biochemistry [1]. DNA containing base analogues has been shown to be modified in thermal stability [2,3,4,5,6], double-helical conformation [7,8] or base-pairing pattern [9,10]. As a result, differences in hybridization strength and selectivity, changes in protein recognition patterns or mutational activity may occur.

2-aminoadenine (2,6-diaminopurine, D) is an adenine analogue which is naturally found in S-2L cyanophage DNA [6]. Replacement of adenine by 2-aminoadenine creates a minor groove modification by introduction of a C2 amino group. This increases the stability of the Watson-Crick base pair with thymine by an additional hydrogen bond (Figure 1). Thermodynamic studies have shown that substitution of 2'-deoxyadenosine by 2-amino-2'-deoxyadenosine stabilized the DNA duplexes significantly [8,11,12]. The introduction of an extra amino group in the minor groove removes the hydration spine which contributes to the stabilisation of the B-DNA structure [13]. It has been observed by spectroscopic studies that 2-aminoadenine may facilitate a B -> Z transition [7,8,14] or even B -> A transition in DNA [15].

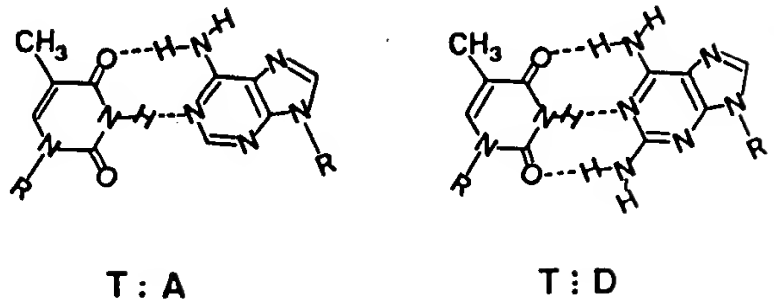


FIGURE 1

Modified bases can be incorporated into DNA by chemical or enzymatic procedures. We have developed a chemical method to introduce 2-aminoadenine into oligodeoxyribonucleotides using protected nucleosides and standard phosphoramidite or phosphotriester methodologies [11]. We have chemically prepared the 2-amino-2'-deoxyadenosine-5'-triphosphate (dDTP) and have shown that dDTP is a substrate for in vitro DNA polymerase I (Klenow fragment) mediated enzymatic synthesis. We show that DNA containing 2-aminoadenine are good substrates for the DNA modifying enzymes T4 DNA ligase, polynucleotide kinase and E.coli methylase. The combined use of these chemical and enzymatic methods allows the preparation of DNA containing 2-aminoadenine at defined sites on one or both DNA strands.

Oligonucleotide probes of defined sequence are important tools for the identification and isolation of specific DNA, cDNA or RNA sequences [16]. High specificity during the hybridization process is the most important feature and the basis for the widespread use of synthetic oligonucleotide probes. Only perfectly matched probe-target DNA duplexes are formed under controlled hybridization conditions [17]. The introduction of additional stabilization to DNA duplexes would allow one to use more stringent hybridization conditions and therefore increase the probe specificity for its target. This should facilitate the screening and analysis of highly complex DNA libraries. We report the use of 2-aminoadenine oligodeoxynucleotides as hybridization probes to screen phage or genomic DNA by Southern blotting.

The effects of base analogues on the cleavage by restriction endonucleases of DNA [18-20] or oligodeoxyribonucleotides [21-24] have contributed to a better understanding of the mode of action of restriction enzymes or methylases [25,26]. Recognition of a specific DNA sequence by an enzyme is achieved by the formation of hydrogen bonds between specific amino acids side chains of the protein and determinants in the major and minor

grooves of DNA. Alteration of bases in a recognition site sequence may therefore disrupt or modify the hydrogen bonding pattern by electronic or steric effects. This usually results in a loss or decrease of the enzymatic activity. We have used 2-aminoadenine-substituted DNA to probe minor groove determinants and have demonstrated that hemi- or full-substitution of adenine by 2-aminoadenine in DNA has profound effects on the recognition and cleavage by some restriction endonucleases.

MATERIALS AND METHODS

Enzymes and Reagents

T4 polynucleotide kinase was kindly donated by N. Murray (University of Edinburgh). Deoxynucleoside triphosphates and E. coli DNA polymerase I (Klenow fragment) were purchased from Boehringer Mannheim GmbH. Restriction endonucleases, EcoRI methylase and T4 DNA ligase were from New England Biolabs, Inc., Beverly, MA. S-Adenosyl-L-[methyl-^{14}C] methionine (0.5 mCi/mmol) was obtained from Amersham Inc. plc, Little Chalfont, Buckinghamshire, England. [γ-^{32}P]-ATP (6,000 Ci/mmol) and GeneScreenPlus membranes were obtained from DuPont-NEN, Boston, MA.

Synthesis of Oligonucleotides

Oligodeoxyribonucleotides were synthesized using the solid-phase phosphoramidite method on an automated machine (Applied Biosystems Model 380A) as described earlier [11].

Synthesis of 2-amino-2'-deoxyadenosine-5'-triphosphate (dDTP)

2-amino-2'-deoxyadenosine [11] was converted to its 5'-monophosphate derivative (dDMP) in 30% yield by the procedure of Kohler et al. [27]. Data for dDMP are : UV (H_2O): 255 (7,600), 278 (8,400); MS (FAB): 347 $[M+H]^+$; TLC (polyethyleneimine-coated cellulose plates-Merck; 0.1 M Na-phosphate pH 6.8/$(NH_4)_2SO_4$/EtOH 100:60:2) R_f 0.10; descending paper chromatography (Whatman #1) (isobutyric acid/NH_4OH/H_2O 66:1:33) R_f 0.62.

dDMP was then converted to its corresponding triphosphate (dDTP) according to the procedure reported by Kahn et al. [28]. UV: as for dDMP. The purification was achieved by descending paper chromatography as above, R_f for dDTP 0.32. All D derivatives give fluorescent spots under UV light.

Enzymatic reactions with oligodeoxyribonucleotides

Oligodeoxyribonucleotides were labelled at their 5'-termini using [γ-^{32}P]-ATP (6,000 Ci/mmol) and T4 polynucleotide kinase as described [29].

Reannealing: The ^{32}P-labelled oligonucleotides (DAP-1; DAP-2; A-3 or DAP-3 and A-4 or DAP-4; DAP-5, DAP-6, DAP-7 and DAP-8) were reannealed by heating at a concentration of 1 µM in a buffer 50 mM TrisHCl pH 7.6, 10 mM

$MgCl_2$ and 50 mM NaCl for 2 min at 95°C and slow cooling to 4°C over 1-2 h.

DNA polymerase reactions: The sample from the reannealing was made up to a final concentration of 5 mM dithiothreitol, 0.05 mM each deoxynucleoside triphosphate (except dDTP which was 0.5 mM) and 100 U/ml E. coli DNA polymerase I (Klenow fragment), incubated for 1 h at room temperature and terminated by heating at 95°C for 2 min. After each enzymatic reaction the oligonucleotides were ethanol precipitated and purified by denaturing polyacrylamide/7M urea, TBE (90 mM Tris borate pH 8.3, 4 mM EDTA) gel electrophoresis [30]. The gels were autoradiographed and the bands containing the desired oligonucleotides were excised. The oligonucleotides were recovered by electroelution in 0.2 x TBE.

Ligase reactions: To the reannealing reaction, dithiothreitol, ATP and T4 DNA ligase were added, respectively to final concentrations of 5 mM, 70 µM and 20 U/ml and the reaction was incubated at 4°C for 12-18 h. The oligonucleotides DAP-1, DAP-2, A-5,6,7,8 and DAP-5,6,7,8 were ligated in this manner.

Denaturing temperature of oligonucleotides containing 2-aminoadenine

Complementary oligonucleotides (60 µM in 0.1 M NaCl, 0.01 M sodium cacodylate pH 7.0) were mixed in equimolar quantities, reannealed by heating at 95°C for 2 min and slow cooling to 4°C over 1 h. Absorbance at 260 nm was recorded as a function of the temperature between 20°C and 80°C at a rate of 1°C/min on a Hewlett-Packard 8450A spectrophotometer. For the denaturation of hybrids to target plasmid DNA spotted on nitrocellulose filters, the method described earlier was used [31].

Hybridization

Southern blots of human genomic B-cell DNA (20 µg/lane) and lambda-DNA containing a 3 kb EcoRI insert with the human TNF gene [32] (0.05 µg to 2.7 µg made up to 5 µg/lane with rabbit genomic liver DNA) were digested with the restriction nuclease EcoRI, separated on an agarose gel, blotted onto GeneScreenPlus (DuPont-NEN) nylon membranes according to Southern [33]. The DNA was crosslinked to the membrane by UV irradiation as reported [34]. Southern blots were prehybridized in 5x SSC, 20 mM sodium phosphate pH 7.0, 10x Denhardt's (0.2% BSA, 0.2% polyvinylpyrrolidone, 0.2% Ficoll, 0.2% SDS), 100 µg/ml sonicated salmon sperm DNA, 7% SDS at 45°, 50° or 62°C for at least 1 h. Hybridization was performed with 0.5 pmole probe/ml in prehybridization buffer at 45°, 50° or 62°C respectively. All probes were radiolabelled with [γ-^{32}P] ATP and T4 polynucleotide kinase at a specific activity of 1.4-1.6x10^7 cpm/pmole. Filters were first washed at the hybridization temperature in 0.2% polyvinylpyrrolidone, 0.2% Ficoll-400,

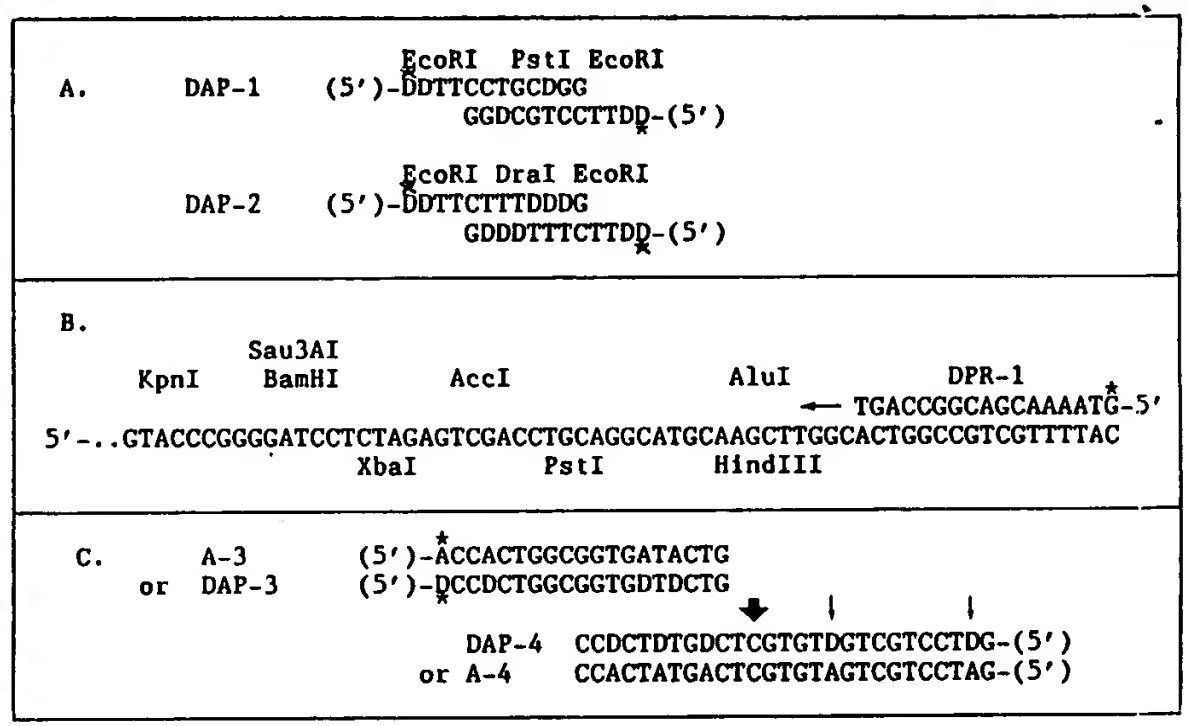


FIGURE 2: DNA substrates used in enzymatic reactions. See text. Stars indicate positions of ^{32}P labels.

0.2% BSA, 0.05 M Tris HCl pH 7.5, 1M NaCl, 0.1% sodium pyrophosphate, 1% SDS for twice 30 min, then washed in either 6x SSC, 0.1% SDS (high salt) or 3.2 M tetramethylammonium chloride (TMAC), 1% SDS [35] at increasing temperatures (2-3°C increments). Hybridization was detected by autoradiography.

Restriction endonuclease reactions

Restriction enzyme digestions were essentially performed as described [36] using the buffers recommended by the suppliers. Aliquots of 0.5 pmol control DNA or D-substituted DNA (5,000-10,000 cpm) in 10 µl were incubated at 37°C for 90 min with various amounts of enzymes (0.5 to 30 U). The reactions were terminated by addition of 10 µl 0.1 M EDTA. The resulted DNA fragments were analyzed by 10% polyacrylamide/7 M urea gel electrophoresis followed by autoradiography.

Methylase reaction

The self-ligated DAP-2 DNA (Figure 2) and control A-2 (A instead of D) respectively (900 pmole) in 30 µl 100 mM TrisHCl pH 8.0, 10 mM EDTA, 0.1 mM S-adenosyl-L-[methyl-^{14}C]methionine (0.5 mCi/mmol) was incubated with 100 U of EcoRI methylase at 37°C for 2 h. The reaction was terminated by heating

at 70°C for 10 min, the DNA was ethanol precipitated and counted in a scintillation counter.

RESULTS

Denaturing temperatures of oligonucleotide probes duplexes containing 2-aminoadenine

We determined the denaturing temperatures of pentadecamer duplexes spectrophotometrically at 260 nm (Table 1). All denaturation curves had a sigmoidal shape. The increase in absorbance (20-25% upon complete denaturation) indicated a normal melting transition pattern. The average effect on the T_m was an additional stabilization of 0.5-1.0°C for each substitution of adenine by 2-aminoadenine. Using the filter hybridization method we observed that introduction of two DC mismatches in a pentadecamer probe (Table 2) caused a 13°C decrease in T_m during hybridization to a linearized plasmid DNA target containing the complementary sequence. With five DC mismatches no hybridization occured 20°C below the T_m of the corresponding reference oligonucleotides without mismatches.

Southern blot hybridization studies with probes containing 2-aminoadenine

DNA-DNA hybridization studies were carried out using bacteriophage λ DNA containing an insert with the human TNF gene [32] or human genomic DNA. Two sets of 20-mer probes for the human TNF gene (I and II, Table 3) were used. Each set is composed of two oligonucleotides differing only by the substitution of D for A. These probes were selected for their high A (or D) content. In addition, a third GC-rich 20-mer probe (PA-3) was used as reference. λ DNA was hybridized at 45°C with either the A-containing probe PA-1, the D-containing probe PD-1 or PA-3 followed by washes in high salt buffer (6x SSC). PA-1, PD-1 and PA-3 were washed off at 61°C, 67°C and 72°C respectively. This indicated an increase in binding strength for the 2-aminoadenine probe of 6°C for 8 DT base pairs. When the hybridization was done at 50°C and the filter directly washed at 62°C we only detected binding of PD-1 and PA-3. Similarly, when the hybridization and washes were at 62°C only PD-1 and PA-3 were found to bind. However, when selection was based on probe length rather than base content (3.2 M TMAC) [35], all three probes PA-1, PD-1 and PA-3 were washed off at 63°C.

We obtained the same results in hybridization to Southern blots of human genomic DNA with the set of probes II and PA-3. However, the background due to non-specific binding was always high for both D-probes PD-1 and PD-2 on genomic DNA targets and washing of the membranes at 95°C failed to completely remove the background radioactivity bound to the DNA.

TABLE 1: Melting temperatures of pentadecamers duplexes in 0.1 M NaCl, pH 7.0.

Duplex	T_m (°C) (± 0.5°)	Number of DT base pairs
1. CTGAAACCGGCGAAG GACTTTGGCCGCTTC	60	0
2. CTGAADCCGGCGDAG GACTTTGGCCGCTTC	61.5	2
3. CTGADDCCGGCGAAG GACTTTGGCCGCTTC	62	2
4. CTGDDDCCGGCGAAG GACTTTGGCCGCTTC	62	3
5. CTGDDDCCGGCGDDG GACTTTGGCCGCTTC	63.5	5

Preparation and properties of DNA containing 2-aminoadenine

Using a N-protected 2-amino-2′-deoxyadenosine phosphoramidite derivative as a building block for the introduction of D, oligodeoxynucleotides from 12 to 40 bases, containing up to nine D bases were synthesized, isolated, purified and analysed as described earlier [11]. Oligonucleotides having a 5′-terminal D nucleoside can be efficiently labelled with [γ^{32}P]-ATP and T4 polynucleotide kinase [11,29]. Kination rates and efficiencies are comparable to normal bases. The oligonucleotide linkers DAP-1 and DAP-2 (Figure 2, panel A), self-complementary at their 3′-termini, were oligomerized by ligation (> 1000 bp on polyacrylamide gels,

Oligonucleotide	T_m °C	Number of DT base pairs	Number of DC mismatches
CTGDDDCCGGCGDDG	49	5	0
CTDDDDCCGGCDDDG	36	5	2
CTDAAACCDDCDAAG	< 29	0	4
CTDDDDCCDDCDDDG	< 29	5	4

TABLE 2: Melting temperatures of pentadecaoligonucleotides - plasmid DNA duplexes. Underlined bases make DC mismatches. The plasmid target contains the complementary sequence (3′) GACTTTGGCCGCTTC (5′). The hybridization temperature was 29°C.

TABLE 3: Oligonucleotide probes to human TNF used in hybridization experiments and number of D bases.

I.	PA-1	GCAATGATCCCAAAGTAGAC	0
	PD-1	GCDDTGDTCCCDDDGTDGDC	8
II.	PA-2	ACGATCAGGAAGGAGAAGAG	0
	PD-2	DCGDTCDGGDDGGDGDDGDG	9
	PA-3	CTGGAGCCCTGGGGCCCCCC	0

data not shown) and used in the study of restriction enzyme cleavages.

Enzymatic synthesis of DNA using dDTP and E. coli polymerase (Klenov fragment) was accomplished with the oligonucleotides A-3/A-4 (Figure 2, panel C) and different mixtures of deoxynucleoside triphosphates (0.05mM). Polymerization in presence of dDTP, dCTP, dGTP and dTTP resulted in incorporation of D on T template with a rate of 20% in comparison with the control reaction (dATP). Similar observations have already been reported [37]. At 0.5 mM dDTP the incorporation was increased to 90% of control. As expected, when dATP, dCTP and dTTP were used without dGTP in the fill-out polymerization the reaction stops at the first C encountered on the template (large arrow in Figure 2, panel C). However, when, in the same experiment, a dDTP, dCTP, dTTP mixture was used, the polymerization did not completely stops at the first C of the template but partially read through by incorporating D on this C. Conversely, polymerization with Klenov polymerase with the fragments A-3/DAP-4 (Figure 2, panel C) in presence of the four nucleotide triphosphates of A,C,G,T or D,C,G,T gave the double-stranded fragment. Omission of dTTP yielded a mixture of shorter fragments (little arrovs in Figure 2, panel C) and full length fragment having incorporated C on a D template. Sequences were confirmed by modified chemical degradation sequencing [38] on the polymerase-elongated oligonucleotides A-3 and DAP-3. The base D behaves like A and gives a positive signal only for the purine reaction (formic acid). Enzymatic primer extension synthesis using dDTP, dCTP, dGTP and dTTP with DNA polymerase I (Klenov fragment) was performed on the M13mp18 single-stranded DNA template using the 17-mer primer DPR-1 (Figure 2, panel B) and generated a double-stranded DNA fragment containing the M13mp18 polylinker restriction sites "hemi-substituted" by D.

Restriction enzyme cleavages

Oligomeric $(DAP\text{-}1)_n$ and $(DAP\text{-}2)_n$ substrates containing D (Figure 2,

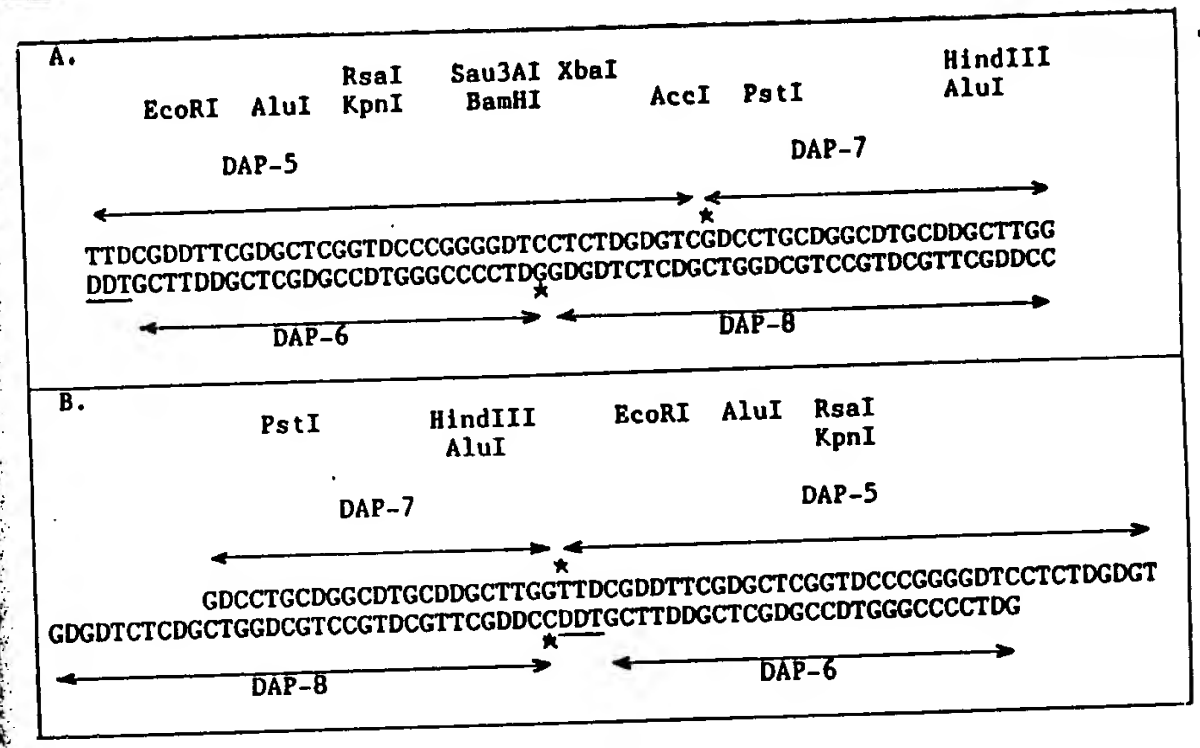


FIGURE 3: Reconstructed pUC18 polylinker DNA containing 2-aminoadenine obtained by reannealing, ligation and polymerase fillout (underlined bases) of the four oligonucleotides DAP-5,6,7,8. Identical reference fragments were prepared by the same method using A-containing oligonucleotides. Restriction enzymes recognition sites are outlined. * indicates position of ^{32}P label.

panel A) were cleaved by EcoRI (GDDTTC) to the monomeric 12-mers. The cleavage was complete after 15 h compared to 1h for the control oligomers containing A. In addition $(DAP\text{-}1)_n$ was also cleaved by PstI (CTGCDG). In contrast, $(DAP\text{-}2)_n$ was not cleaved by DraI (TTTDDD). By radiolabelling experiment ve found that ^{14}C-methyl groups were transferred to $(DAP\text{-}2)_n$ by the EcoRI methylase [39] when S-adenosyl-L-[methyl-^{14}C]methionine was used as methylation agent, and conferred resistance to EcoRI cleavage.

Cleavage was also tested on hemi-substituted DNA (Figure 2, panel B). The results are summarized in Table 4. Substitution of A by D in one strand of DNA completely inhibited XbaI digestion and sloved down the rate of cleavage by the enzymes AccI, AluI, BamHI, HinfI, RsaI and Sau3AI. There was no noticeable change for cleavage by EcoRI, HindIII, KpnI and PstI.

Finally, a fully substituted DNA substrate (Figure 3) was treated by restriction enzymes (Table 4). Only PstI, and at a slover rate KpnI and AluI, cleaved the DNA. Cleavage was fully inhibited for the enzymes AccI, BamHI, HindIII, HinfI, RsaI, Sau3AI, XbaI and, surprisingly, EcoRI. The failure of EcoRI to cleave contrasts with the results with $(DAP\text{-}1)_n$ and

Subst.	AccI	AluI	BamHI	EcoRI	HindIII	HinfI	KpnI	PstI	RsaI	Sau3AI	XbaI
Site	GTCGAC	AGCT	GGATCC	GAATTC	AAGCTT	GANTC	GGTACC	CTGCAG	GTAC	GATC	TCTAGA
HEMI	VS	S	VS	1	1	S	1	1	VS	S	0
FULL	0	VS	0	S	0	0	S	1	0	0	0

TABLE 4: Cleavage of DNA substrates containing the analogue 2-aminoadenine by restriction enzymes. The hemi-substituted DNA substrate is the in vitro DNA polymerase I (Klenow) primer extension product depicted in Figure 2, panel B. The fully substituted DNA substrates are described in Figure 3. Rates of cleavage are: 1 = > 50% of A-containing control DNA; S = slow (10-50% of control); VS = very slow (up to 10% of control); 0 = no cleavage detectable.

$(DAP-2)_n$ DNA. Attempts to detect EcoRI cleavage after 20 h at 37°C or 24 h at 16°C (to minimize DNA breathing) were unsuccessful. However, we found that EcoRI could cleave, at a slower rate, the substrate if the EcoRI site was closer to the center (Figure 3, panel B). As above, PstI, KpnI and AluI cleaved whereas HindIII and RsaI did not cleave.

DISCUSSION

The combined use of oligodeoxyribonucleotide chemical synthesis and enzymatic techniques allows the introduction of the analogue D with selectivity and flexibility at the desired positions on one or both strands of duplex DNA. A nucleoside D at the 5'-terminus can be phosphorylated with polynucleotide kinase and joined to another fragment with T4 DNA ligase. DNA containing D can be used as a template for DNA polymerase I (Klenow fragment) - mediated synthesis of the opposite strand with incorporation of T on template D. Induction of mutations through the formation of DC mispairs has been reported [10] but the frequency of these mutations, while higher than spontaneous mutations, is insignificant [40].

Nucleoside base-pair stacking and hydrogen bonds are the main factors controlling duplex stability. The DT base pair with three hydrogen bonds in a Watson-Crick arrangement is comparable to GC although less stable. This suggests differences in the base pair stacking pattern and possibly in the geometry of the hydrogen bonds. We demonstrate that, in high ionic strength salt buffer, probes containing D are more specific than their normal counterparts. We can show selective hybridization of only the D-containing probe to a specific target on bacteriophage λ DNA. As expected, posthybrization washes in tetramethylammonium chloride [35] eliminated all

base composition effects. Pools of oligonucleotides representing all possible coding sequences for a region of amino acid sequence are commonly used to screen libraries for the identification of cloned DNA. A potentially useful application would be to substitute all adenines at non-redundant sites by 2-aminoadenine and hybridize at higher stringency. Hybridisation of probes containing D to genomic DNA gave some non-specific background which might constitute a caveat to the use of D for probing genomic DNA.

A crystal structure of the complex between the EcoRI endonuclease and a short oligonucleotide has been recently determined and has accurately localized in the major groove the DNA determinants important in sequence recognition [41]. As a result we would anticipate a minor, if any, perturbation of EcoRI cleavage on DNA containing the analogue D (GDDTTC). This is what we observed on the substrates $(DAP-1)_n$ and $(DAP-2)_n$. The retardation on the rate of cleavage with EcoRI can tentatively be attributed to a conformational constraint preventing the formation of Type I neokink necessary for protein binding [41]. Our results support the work of Gumport et al. on singly substituted oligonucleotides [24,25].

Hemi-substituted DNA substrates, in which only one strand of the duplex is substituted, have already been used in similar studies with nucleotide analogues including 2-aminoadenine [22,37]. Effects observed on such substrates may however be very different than effects observed on substrates having sites substituted on both strands [42]. Our assay on hemi-substituted substrates can only detect cleavage in the analogue-containing strand because the label resides on the polymerase primer (Figure 2, panel B). Results from Table 4 show dramatic differences between hemi- and fully-substituted DNA substrates. Differences in the cleavage by EcoRI of fully substituted 63 bp fragments (Figure 3) may reflect the existence of conformational heterogeneity in D-containing DNA and its importance to the protein recognition mechanism.

All fragments obtained from cleavage of D-containing DNA were similar to reference natural sequence DNA fragments. Cleavage of DNA by restriction enzymes is accomplished in two steps: 1) binding of the enzyme to the recognition site and 2) cleavage of a phosphodiester bond. Base analogues may interfere with both of these events. Our assay does not allow to distinguish between these two steps. Nevertheless we are able to assess qualitatively the effect of minor groove modification due to D and to demonstrate complete inhibition cleavage for several restriction endonucleases.

ACKNOVLEDGEMENTS

We thank Angelika Chollet, Jean-Pierre Voegeli and Edgar Ayala for technical assistance.

+Present address: Glaxo Institute for Molecular Biology SA, PO Box, CH-1211 Geneve 24, Switzerland

*To whom correspondence should be addressed

REFERENCES

1. Goeddel, D.V., Yansura, D.G. and Caruthers, M.H. (1978) Proc. Nat. Acad. Sci. USA 75, 3578-3582.
2. Kawase, Y., Iwai, S., Inoue, H., Miura, K. and Ohtsuka, E. (1986) Nucleic Acids Res. 14, 7727-7736.
3. Eritja, R., Horowitz, D.M., Walker, P.A., Ziehler-Martin, J.P., Boosalis, M.S., Goodman, M.F., Itakura, K. and Kaplan, B.E. (1986) Nucleic Acids Res. 14, 8135-8153.
4. Eritja, R., Kaplan, B.E., Mahskar, D., Sowers, L.C., Petruska, J. and Goodman, M.F. (1986) Nucleic Acids Res. 14, 5869-5884.
5. Seela, F. and Driller, H. (1985) Nucleic Acids Res. 13, 911-926.
6. Kirnos, M.D., Khudyakov, I.Y., Alexandrushkina, N.I. and Vanyushin, B.F. (1977) Nature 270, 369-370.
7. Taboury, J.A., Adam, S., Taillandier, E., Neumann, J.-M., Tran-Dinh, S., Huynh-Dinh, T., Langlois, B., Conti, M. and Igolen, J. (1984) Nucleic Acids Res. 12, 6291-6305.
8. Gaffney, B.L., Marky, L.A. and Jones, R.A. (1984) Tetrahedron 40, 3-13.
9. Hopkins, R. and Goodman, M.F. (1980) Proc. Nat. Acad. Sci. USA 77, 1801-1805.
10. Poltev, V.I., Shulyupina, N.V. and Bruskov, V.I. (1977), Molekulyarnaya Biologiya 13, 822-828.
11. Chollet, A., Chollet-Damerius, A. and Kawashima, E.H. (1986) Chemica Scripta 26, 37-40.
12. Brennan, C.A. and Gumport, R.I. (1985) Nucleic Acids Res. 13, 8665-8684.
13. Kopka, M.L., Fratini, A.V., Drew, H.R. and Dickerson, R.E. (1983) J. Molec. Biol. 163, 129-146.
14. Coll, M., Wang, A.H.J., van der Marel, G.A., van Boom, J.H. and Rich, A. (1986) J. Biomolec. Struct. & Dynamics 4, 157-172.
15. Borah, B., Cohen, J.S., Howard, F.B. and Miles, H.T. (1985) Biochemistry 24, 7456-7462.
16. Suggs, S.V., Wallace, R.B., Hirose, T., Kawashima, E.H. and Itakura, K. (1981) Proc. Natl. Acad. Sci. U.S.A. 11, 6613-6617.
17. Wallace, R.B., Shaffer, J., Murphy, R.F., Bonner, J., Hirose, T. and Itakura, K. (1979) Nucleic Acids Res. 6, 3543-3557.
18. Berkner, K.L. and Folk, W.R. (1979) J. Biol. Chem. 254, 2551-2560.
19. Huang, L.H., Farnet, C.M., Ehrlich, K.C. and Ehrlich, M. (1982) Nucleic Acids Res. 10, 1579-1591.
20. Marchionni, M.A. and Roufa, D.J. (1978) J. Biol. Chem. 253, 9075-9081.
21. Ono, A. and Ueda, T. (1987) 15, 219-232.
22. Ono, A. and Ueda, T. (1987) 15, 3059-3072.
23. Jiricny, J., Wood, S.G., Martin, D. and Ubasava, A. (1986) 14, 6579-6590.
24. Brennan, C.A., Van Cleve, M.D. and Gumport, R.I. (1986) J. Biol. Chem. 16, 7270-7278.
25. Brennan, C.A., Van Cleve, M.D. and Gumport, R.I. (1986) J. Biol. Chem. 16, 7279-7286.
26. Modrich, P. and Rubin, R.A. (1977) J. Biol. Chem. 252, 7273-7278.

27. Kohler, P., Wachtl, M. and Tamm, C. (1980) Helv. Chim. Acta 63, 2488-2494.
28. Kahn, N.N., Wright, G.E., Dudycz, L.W. and Brown, N.C. (1985) Nucleic Acids Res. 13, 6331-6342.
29. Van de Sande, J.H., Loewen, P.C. and Khorana, H.G. (1972) J. Biol. Chem. 247, 6140-6143.
30. Maniatis, T., Jeffrey, A. and van de Sande, J.H. (1975) Biochemistry 14, 3787-3794.
31. Chollet, A. and Kawashima, E. (1985) Nucleic Acids Res. 13, 1529-1541.
32. Marmenout, A., Fransen, L., Tavernier, J., van der Heyden, J., Tizard, R., Kawashima, E., Shaw, A., Johnson, M.J., Semon, D., Müller, R., Ruysschaert, M.R., von Vliet, A. and Fiers, W. (1985) Eur. J. Biochem. 152, 515-522.
33. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
34. Church, G.M. and Gilbert, W. (1984) Proc. Nat. Acad. Sci. U.S.A. 81, 1991-1995.
35. Wood, W.I., Gitschier, J., Lasky, L.A. and Lawn, R.M. (1985) Proc. Nat. Acad. Sci. U.S.A. 82, 1585-1588.
36. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) in Molecular Cloning, A Laboratory Manual, edited by CSH.
37. Bodnar, J.W., Zempsky, W., Warder, D., Bergson, C. and Ward, D.C. (1983) J. Biol. Chem. 258, 15206-15213.
38. Banaszuck, A.M., Deugan, K.V., Sherwood, J., Michalak, M. and Glick, B.R. (1983) Anal. Biochem. 128, 281-286.
39. Rubin, R.A. and Modrich, P. (1977) J. Biol. Chem. 252, 7265-7272.
40. Alderson, T. (1964) Mutation Res. 1, 205-208.
41. McClarin, J.A., Frederick, C.A., Wang, B.C., Greene, P., Boyer, H.W., Grable, J. and Rosenberg, J.M. (1986) Science 234, 1526-1541.
42. Streeck, R.E. (1980) Gene 12, 267-275.

ELSEVIER

Genetic Analysis: Biomolecular Engineering
13 (1996) 15-21

GENETIC ANALYSIS
Biomolecular Engineering

Oligonucleotides containing 2-aminoadenine and 5-methylcytosine are more effective as primers for PCR amplification than their nonmodified counterparts

Y. Lebedev[a,*], N. Akopyants[b], T. Azhikina[a], Y. Shevchenko[a], V. Potapov[a], D. Stecenko[a], D. Berg[b], E. Sverdlov[a]

[a] *Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Miklukho-Maklaya, 16/10, 117871 Moscow , Russia*
[b] *Department of Molecular Microbiology, Box 8230; Washington University School of Medicine, 4566 Scott Ave; St. Louis, MO 63110 , USA*

Received 26 October 1995; revised 15 February 1996; accepted 15 February 1996

Abstract

Oligonucleotides containing the modified bases 5-methylcytosine and 2-aminoadenine in place of cytosine and adenine, respectively, have higher than normal affinity for complementary sequences. The strong binding oligonucleotides (SBO) are much better than their normal counterparts in PCR amplification: they yield significantly more product per cycle, allow amplification at annealing temperatures as high as 72°C and, unlike their normal counterparts, allow efficient priming from within a palindromic sequence. We propose that such strong binding oligonucleotides will be valuable in numerous PCR applications, including: (i) minimization of the frequency of mutants among PCR products; (ii) when only short specific primers can be designed based on available sequence information; (iii) when the material available for the analysis is limited in quantity: and (iv) when primer binding is blocked by DNA secondary structure involving a primer binding site, or chain extension is impeded by secondary structure in downstream sequences.

Keywords: Modified oligonucleotide; PCR primer; PCR amplification

1. Introduction

The polymerase chain reaction (PCR) method has become one of the most powerful and widely used tools in modern biology, having revolutionized areas as diverse as analyses of gene structure, genetic regulatory mechanisms and genome organization, infectious disease and epidemiology, evolution and forensics [1–4]. For all its power, however, further optimization could bring considerable benefits. For example, numerous errors, both base substitutions and internal deletions, accumulate cycle by cycle during amplification, reflecting error-prone DNA priming and synthesis and probably thermal damage to template DNA, as well as selection for templates that are more easily replicated [1–3,5–8]. Reducing the number of cycles needed for a given level of amplification could dramatically lessen the problem of mutant accumulation, which can be especially severe when amplification proceeds from minute quantities of target DNA or RNA, such as from a single cell. Second, PCR amplification of certain regions tends to be quite inefficient, due to hairpin and perhaps other secondary structures in the template strand that overlap or lie downstream of a primer binding site, where they can block primer binding or impede continued DNA synthesis, respectively. Third, in many applications, information about target sequences is very limited, as when primers are designed based on partial amino acid sequences of newly detected proteins, or on evolutionarily conserved motifs. Often the specific primers that can be designed are quite short, and may fail to give efficient specific amplification, especially if they have a high A + T content or too many mismatches. In each of these cases, PCR amplifi-

* Corresponding author. Tel.: +7 095 3306529; Fax: +7 095 3306538; e-mail: sverd@humgen.siobc.ras.ru

PII S1050-3862(96)00139-8

cation would be greatly improved by any change that increased the efficiency of amplification with a given pair of primers. One way to achieve this would be to increase the strength of specific interaction between primer and complement target DNA or RNA strands. The oligonucleotides containing the modified bases 5-methylcytosine (5-MeC) and 2-aminoadenine (2-NH_2A) in place of cytosine and adenine, respectively, have higher than normal affinity for complementary sequences [9,10] (Henceforth these oligonucleotides containing such modified bases will be called "strong binding oligonucleotides" or "SBOs").

Recently we developed and implemented methods for phosphoramidite based automated synthesis of oligonucleotides containing these modified bases and showed that strings of pentanucleotide SBOs can prime DNA sequencing reactions from a complementary sequence under conditions in which the corresponding unmodified oligonucleotides are considerably less effective [11,12]. We also demonstrated that SBO penta- and hexamers prime template directed DNA synthesis more efficiently (Lagutina et al., Genet Anal Biomol Eng, submitted).

Here we show that such SBOs are superior to the corresponding standard oligonucleotides for PCR amplification in terms of rate of accumulation of PCR product per cycle, maximum annealing temperature allowing amplification, and amplification from a primer binding site within palindromic DNA.

2. Materials and methods

2.1. Oligonucleotide synthesis

Modified protected phosphoramidites have been prepared as described in [11,12]. Oligonucleotides were synthesized on a controlled porous glass support with a Milligene 7500 synthesizer. After deprotection by aqueous ammonia at 60°C for 20 h, the oligonucleotides were purified by denaturing polyacrylamide gel electrophoresis. Melting temperatures (T_m) were calculated as in [13] and were increased by N where N is the number of modified residues, in case of modified oligos [11]. Melting temperature (T_m) for modified universal primer was determined in previous experiments [12]. Universal sequencing primer (−20) GTAAAACGACGGCCAGT and "reverse" sequencing primer CAGGAAAACAGCTATGAC, as well as their modified counterparts, containing all C and A substituted by 5-MeC and 2-NH_2A were used for PCR amplification in case of inserts in plasmid pBS, phage M13 and for amplification from within the "frying pan" structure. The primers and their T_m are listed in the Table 1.

2.2. Template DNSs

A clone from heteronuclear cDNA (hncDNA) human chromosome 19 specific library [14] with 284 bp hncDNA insert cloned in EcoRI-BamHI doubly digested M13 mp18 vector was used for ss and ds DNA purification for PCR experiments. Recombinant pBS plasmids containing fragments of *Helicobacter pylori* genomic DNA 1.1 and 2.5 kb in length were used for other PCR experiments.

Fragments of DNA forming "frying pan" structures as shown in Fig. 3 with lengths 528 bp and 312 bp were kindly provided by Dr S. Lukyanov [15,16].

Standard techniques of molecular biology were performed as described in [13].

2.3. PCR amplification

PCR amplification of recombinant M13 mp18 and pBS plasmids were performed at the Perkin Elmer's thermocycler. Template DNA (20 ng) was amplified with 20 pmol each of the primers, 0.5 U Taq. polymerase (Perkin Elmer or samples purified in Dr Berg's lab), 0.25 mM dNTP, 3 mM $MgCl_2$ in the following mode: 95°C, 30 s; 57°C, 30 s; 72°C, 1 min; 5, 10, 15, 20, 25 or 30 cycles. PCR-fragments were analysed by electrophoresis in 2% agarose gel. All the PCR conditions different from those mentioned above are indicated in legends to figures.

2.4. Quantitative PCR

PCR amplification of dsDNA (as well as ssDNA in other non-shown experiments) of recombinant M13 mp18 containing a 284 bp insert, the clone of heterogeneous nuclear cDNA library that was prepared in the our lab [14], were performed with universal sequencing and reverse M13 primers as well as with their modified counterparts as described above with addition of 0.5 pmol α-[^{32}P]dATP (3000 Ci/mmol, Obninsk, Russia). Labelled PCR-fragments were separated in 2% agarose gel and exposed with Kodak XAR-2 film. To determine

Table 1
The structure of primers and their T_m

	Primer	T_m (norm) °C	T_m (mod) °C	ΔT_m °C
Universal	GTAAAACGACGGCCAGT	68[a]	78[a]	10
Reverse	CAGGAAAACAGCTATGAC	52[b]	64[b]	12

[a]T_m were determined experimentally [12].
[b]Calculated T_m

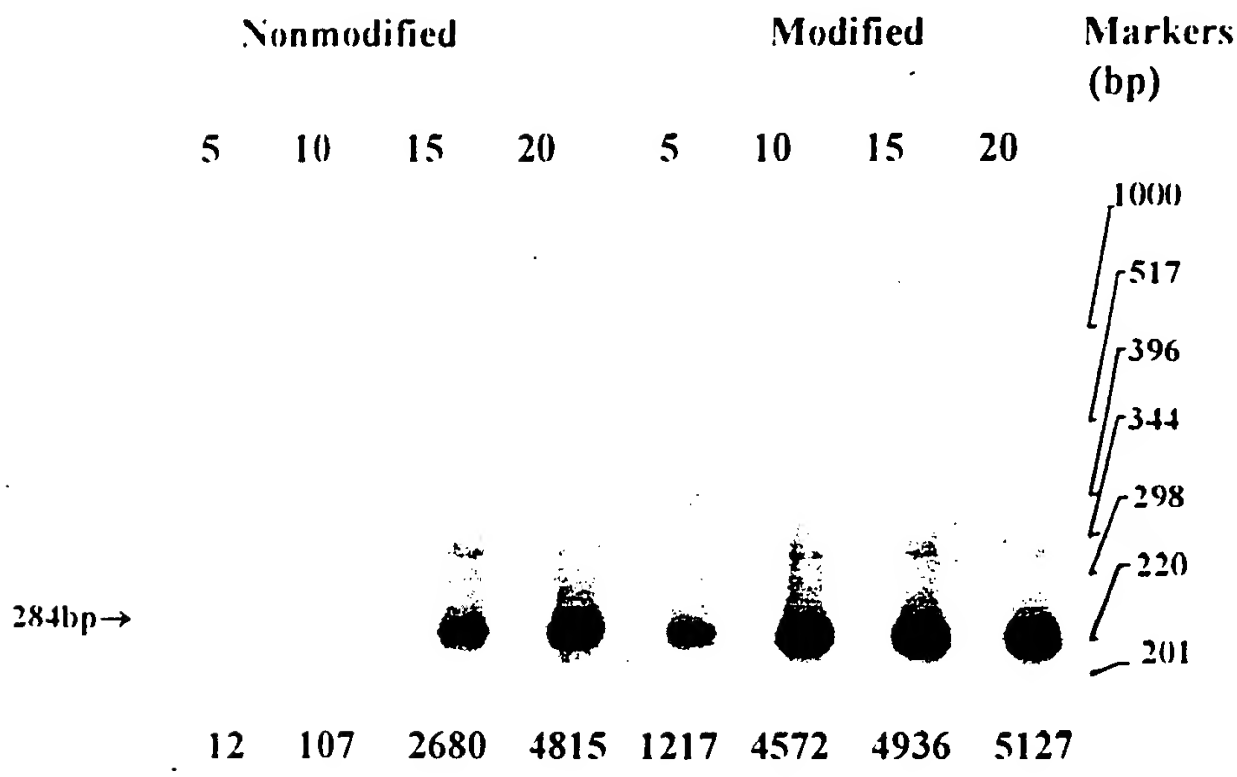


Fig. 1. Influence of the primmer modification on the rate of PCR product accumulation. Radioautograph of 2% agarose gel electrophoresis of products of radioactive PCR amplification of recombinant M13 ds DNA, that is the clone from human chromosome 19-specific heterogeneous nuclear cDNA library is present. PCR amplification with primers GTAAAACGGCCAGT and CAGGAAAACAGCTATGAC or their modified counterparts were carried out after 5, 10, 15 and 20 cycles as indicated. Variants of primers (non-modified or modified) and number of PCR cycles performced are indicated in upper part. Amount of 284bp PCR fragments (designated) was determined as described in Materials and methods and are indicated at the bottom. The marker is 1 kb DNA Ladder (Gibco-BRL).

the relative yields of the PCR products, autoradiographs were subjected to densitometric analysis using the image analyser System 3 (Apple) (program SCAN 1000).

3. Results and discussion

The present experiments were based on findings that (i) replacement of cytosine and adenine with 5-methylcytosine and 2-aminoadenine in DNA can markedly increase the strength of annealing between complementary DNA strands [9,11]; and (ii) the modified oligonucleotides can prime DNA synthesis in a chain termination DNA sequencing reaction more efficiently than their nonmodified counterparts [11].

3.1. Rate of PCR product accumulation.

The increased efficiency of DNA synthesis achieved using modified oligonucleotide primers has important implications for PCR because it allows in principle to increase the yield of the amplification product with the use of the same number of cycles as in case of standard primers, or to reduce the number of cycles without sacrificing final yield. This conclusion directly follows from the simple formula (1): $Y = (X)^n$, where Y represents the final yield of the amplification product after n cycles of PCR provided that the yield after one cycle is equal to X (yield is the quantity of molecules formed per one initial DNA molecule). This formula also demonstrates that the higher the difference in the yields of X between modified and non-modified primers, the stronger is the advantage of the use of the more tightly binding oligos for PCR. For example, an increase in yield per cycle from 1.7 (a rather usual factor with standard oligonucleotides [1–4]) to 1.9 would (i) allow the final yield of product to be increased 5.8×10^5 to 9.3×10^6 in a standard 25 cycle reaction or (ii) cessation of PCR amplification five cycles earlier (between the 20th and 21st cycles) to achieve the yield 5.8×10^5.

Direct tests showed that base modification did indeed increase the efficiency of formation of PCR products with various quantitative value of the differences in yields between modified and non-modified primers. In the case illustrated by Fig. 1, nearly 20 PCR cycles with non-modified primers were required to match the yield attained approximately at the 10th PCR cycle with modified primers. Even if we assume that in the case of modified primers the yield per a cycle (X) equals 2, that is theoretical, then we would have to conclude that the yield X for the standard primers is too low — around 1.56 (formula 1). That low value could be a consequence of some restraints posed by the template on amplification.

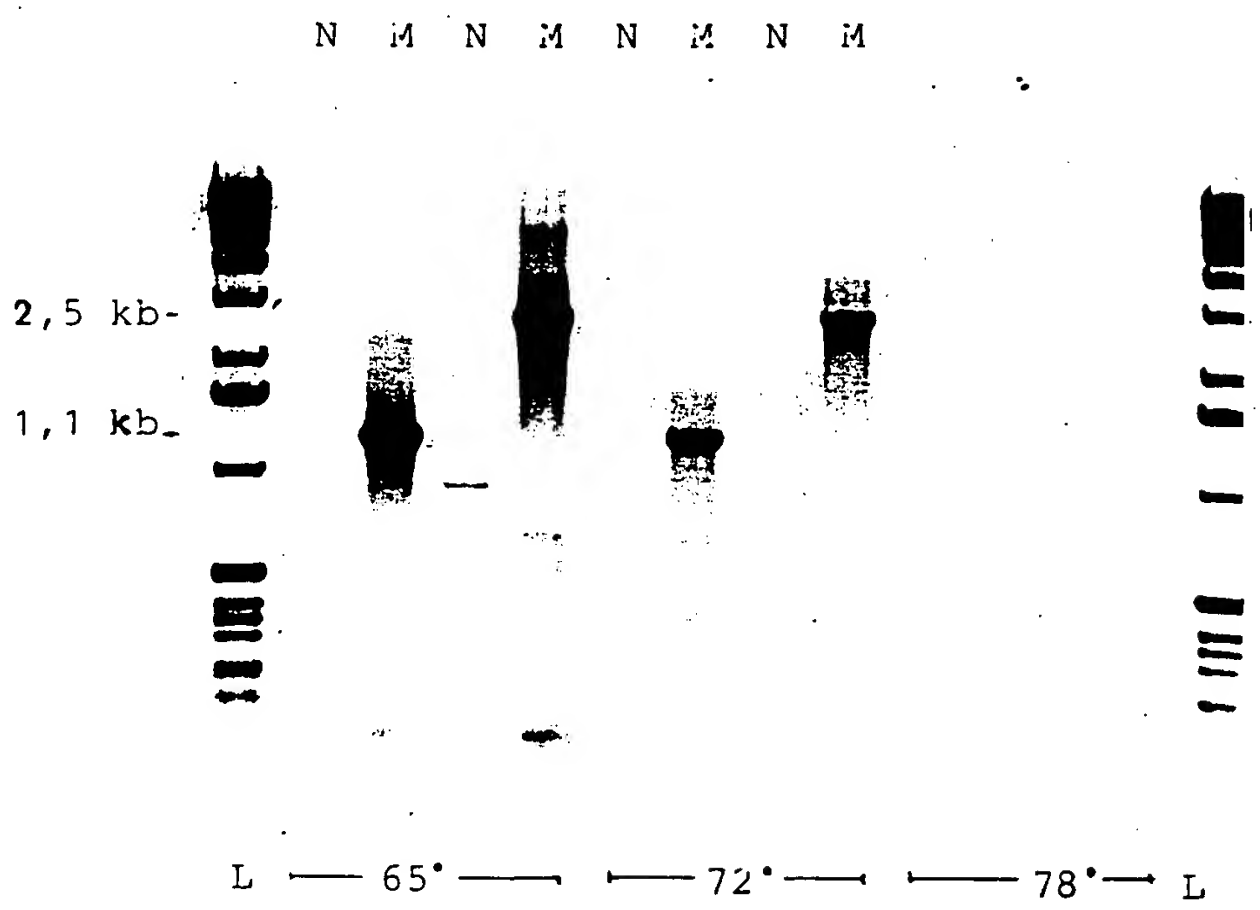


Fig. 2. The temperature dependence of PCR amplicication with modified and normal primers. Oglionucleotiedes GTAAAACGGCCAGT, universal sequencing M13 primer (−20). and CAGGAAAACAGCTATGAC, reverse primer. were used for PCR amplifications: 1.1 kb and 2.5 kb inserts cloned into pBS plasmids were amplified at different annealing temperatures (indicated at the bottom part) and separated by electrophoresis in 1% agarose gel. N, non-modified universal and reverse M13 primers were used for amplification: M, their modified analogous; L. 1 kb DNA Ladder (BRL).

3.2. Annealing temperature effects on PCR amplification.

The increased strength of binding of oligonucleotides with modified bases suggests that they should allow PCR amplification at higher annealing temperatures. Fig. 2 shows an example illustrating that this is indeed the case. Whereas both modified and non-modified primers were effective in PCR amplification at annealing temperatures of 57°C (data are not shown), only the modified primers were effective at an annealing temperature 65°C and more. In addition, the modified primers were almost as effective when the temperature was cycled between 94°C (denaturation) and 72°C (both annealing and extension; 72°C is the temperature typically used for extension during PCR amplification [1–4].

These results show that modified primers are advantageous when high annealing temperatures are needed. We anticipate that the ability conferred by modified oligonucleotides to prime PCR amplification at temperatures equal to those used for extension will be important in cases where hairpins and other secondary structures would form during low temperature annealing steps in PCR with normal primers. Any failure to completely melt such structures before arrival of the 3′ end of the nascent chain would diminish the overall efficiency of PCR amplification and/or provoke deletion and other errors during synthesis.

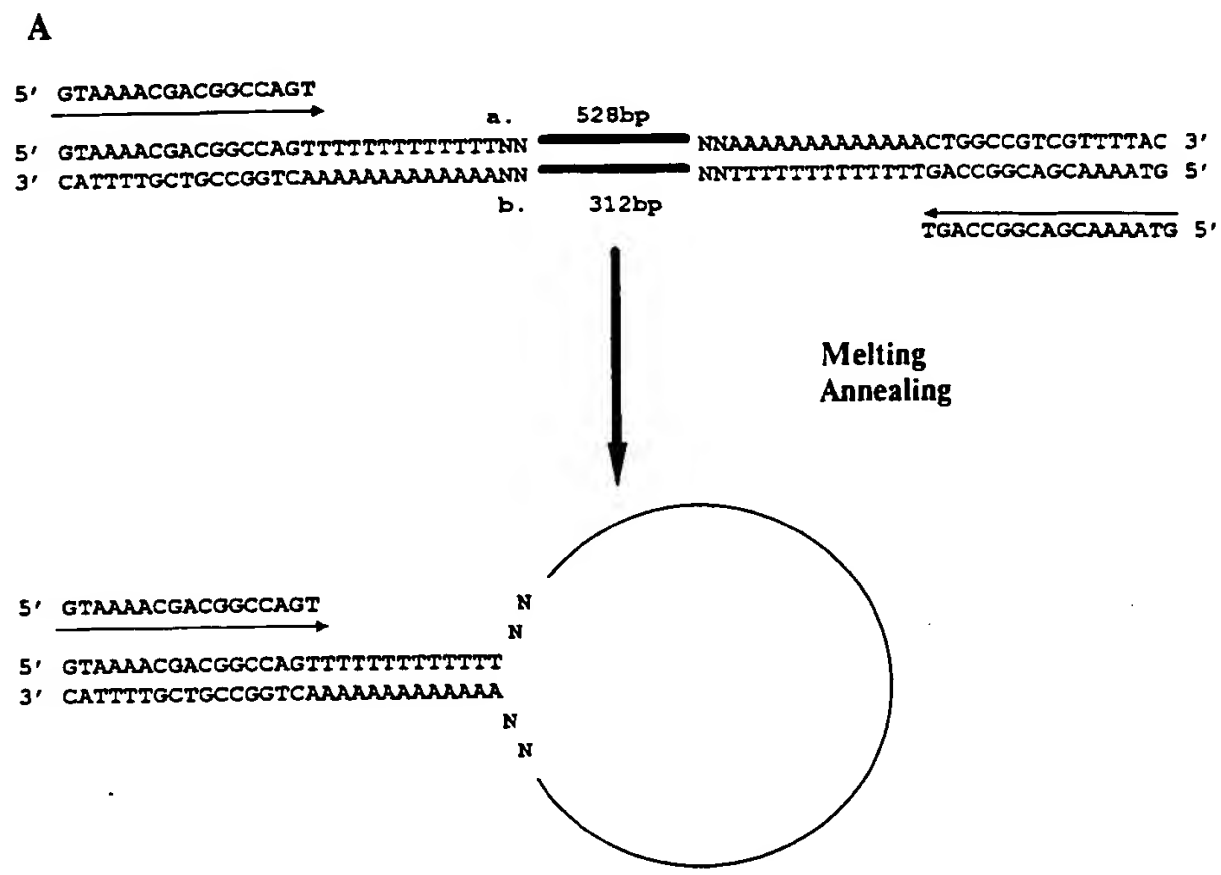
3.3. Modified oligonucleotides overcome effect of primer binding site within hairpin structure.

Special "frying pan" templates (Fig. 3) were used to test whether modified oligonucleotides could prime PCR amplification from primer binding sites within extended hairpin structures. Earlier work [15,16] had shown that such structures strongly suppressed amplification with the natural primers, independent of the temperature used for amplification. Two templates that differed in central regions (between the terminal inverted repeats, which contained primer-binding sites) were used, and in both cases only the modified oligonucleotides proved to be effective as primers for amplification. Evidently, only the modified oligonucleotides are able to displace its normal counterpart in a hairpin snap-back structure, thereby allowing it to prime DNA synthesis, and promote repeated cycles of amplification.

4. Concluding remarks

The results presented here show that oligonucleotides containing base modifications that allow them to bind more strongly to complementary DNA than its normal counterparts should be useful as primers for PCR amplification in several contexts.

Fig. 3
and i
or 25
indic



"Frying pan" structure.

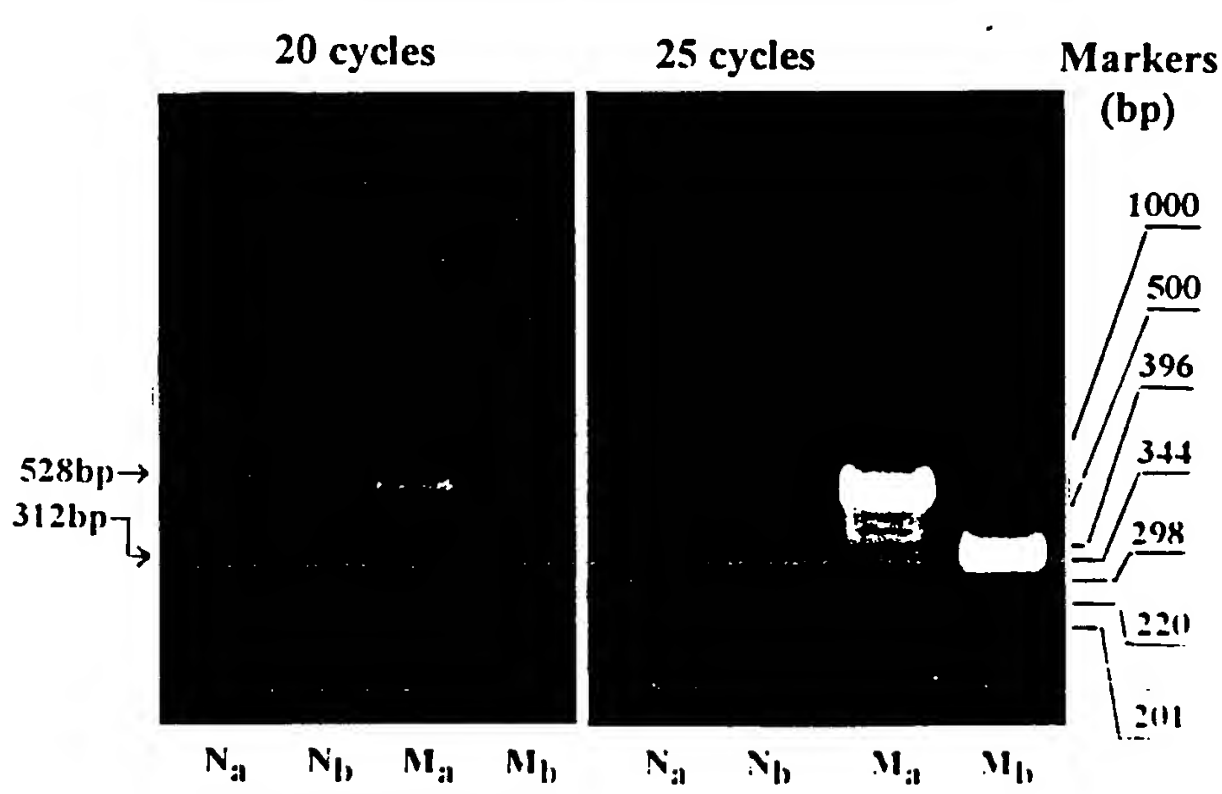


Fig. 3. PCR amplification of DNA fragments forming secondary structure at primer binding sites with universal sequencing M13 primer (−20) and its modified counterpart. (A)A scheme of structure of the template DNA. (B) Agarose gel (2%) with PCR fragments obtained after 20 (left) or 25 (right) cycles of amplification. N and M, results obtained with non-modified or modified primers, respectively. Sizes of PCR fragments are indicated.

(i) When the quantity of DNA or RNA to be amplified is restricted: during amplification from just one or a few cells, detection of pathogens present at low population densities in the environment or during chronic or latent infection; and in forensic medicine.

(ii) When the reduction of the number of PCR cycles is highly desirable. This point needs some special comments. There are several sorts of artifacts taking place during PCR amplification, all of them accumulated with the increase of the number of cycles. The accumulation of the point mutations is the best known example of the mistakes. In general, speaking of accumulation of the mutations, authors discuss the misincorporations (10^{-4}–10^{-5} nucleotides per cycle) and small deletions in the growing DNA chains due to infidelity of the templates copying by the DNA-polymerases [1–4]. In addition, another source of the mistakes can be chemical modification of the template due to rather rigorous conditions of the reaction during the denaturation and high temperature primer extension. If one uses the usual PCR profile (1–2 min at >94°C for denaturation, 1–2 min at 50–55°C for primer annealing and 1–2 min at 72°C) for the extension step and repeats it 30 times, the total time of the incubation of the mixture at the elevated temperature is rather long and one can expect the DNA damages, mainly due to the depurinisation [5,17]. These lesions can cause additional artefacts during the following replication stages — mainly short deletions. The appearance of the sequence changes, even though not very outnumbered by itself, can be further enhanced by biased amplification of those that are more easily replicated.

The biased amplification of the shorter fraction of the restriction fragments has been observed [6] when the whole genome amplification [18–20] has been carried out.

Such a biased amplification depending on the length and on base composition has also been observed in case of two alleles amplification with identical primers [7]. Similar results were described when amplifying several genomic fragments with the multiple pairs of primers [8]. One can expect the same effect when amplifying the complex mixture of the fragments as it takes place for example in case of amplification of chromosomal fragments obtained by microdissection [21,22], amplifications of the cDNA libraries [23] or a single sperm typing with allele-specific oligomers [24]. Another sort of artifacts is DNA recombination during PCR [25,26] leading to chimerical DNA molecules, which could be considerably enhanced in archaeological and forensic application [27]. The increased efficiency of DNA synthesis and thus decrease in required number of cycles made possible by strong binding oligonucleotides should decrease the accumulation of such artifacts.

(iii) When PCR amplification would be improved by use of two-step procedure as described [3] but at higher temperature for primer annealing and extension. We anticipate this will often be the case when the segment to be amplified contains palindromes and other sequence arrangements likely to yield secondary structures during lower temperature annealing steps. Amplification using the same high temperature for annealing and chain extension (72°C) would maximize the melting of such secondary structures, and might decrease the yield of deletion variants, and improve yield as well.

(iv) When the primer binding site is within a potential hairpin or other secondary structure that would impede binding of normal primers. The stronger binding of the modified primer at even normal annealing temperatures apparently allows displacement of the complementary blocking strand, and thus effective initiation of replication. Even though computer programs for improved primer design are under constant development, even the best of such programs can not anticipate inverted repeats, one member of which is in an unsequenced nearby region. The use of strong binding oligonucleotides as primers for PCR, and also for sequencing by primer walking, should allow us to avoid this inherent weakness in current strategies.

At present we have developed and tested just two modifications, 5-methhylcytosine and 2-aminoadenine. Modifications that similarly increase the affinity of thymidine for its complementary adenosine are also feasible, and are being studied by our group. Once the T-modifications that are both easy to construct as phosphoramidites, and effective in increasing affinity for complementary sequences have been identified, we can expect to be able to construct even more potent strong binding oligonucleotides for use as hyper-efficient PCR and sequencing primers and hybridization markers.

Acknowledgements

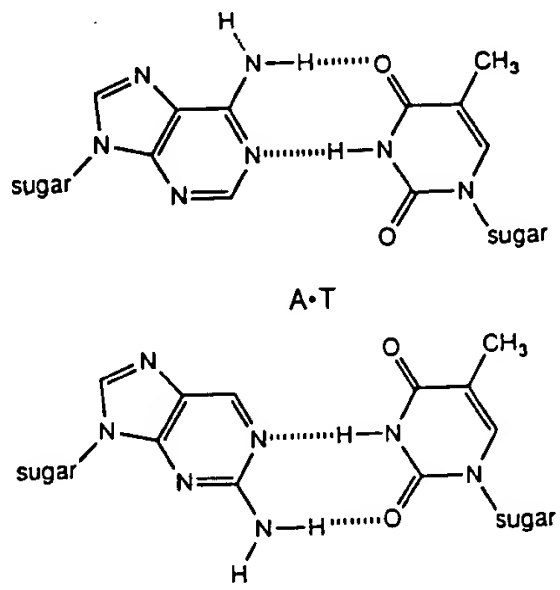
We are grateful to Dr S. Lukyanov for providing the "frying pan" templates. The work was supported by a grant from the Russian Human Genome National Program, Russian Foundation for Fundamental Investigation grant 93-94-7810 and partially supported by NIH grants HG00053, DK48029 and TWO039.

References

[1] Erlich HA, Gelfand D, Sninsky JJ. Science 1991; 252: 1643–1650.
[2] Bloch W. Biochemistry 1991; 30: 2735–2747.
[3] Cha RS, Thilly WG. PCR Methods Appl 1993; 3: 518–529.
[4] Innis MA (editor). PCR Protocols: A Guide to Methods and Applications. CA, Academic Press, 1990.
[5] Douglas A, Atchison B. PCR Methods Appl 1993; 3: 133–134.

[6] Lisitsyn N, Lisitsyn N, Wigler M. Science 1993; 259: 946–951.
[7] Horn GT, Richards B, Klinger KW. Nucleic Acids Res 1989; 17: 2140–2145.
[8] Chamberlain JS, Gibbs RA, Ranier JE, Nguyen PN, Caskey CT. Nucleic Acids Res 1988; 16: 11441–11456.
[9] Hoheisel JD, Craig AG, Lehrach H. J Biol Chem 1990; 265: 16656–16660.
[10] Cheong C, Tinoco I Jr., Chollett A. Nucleic Acids Res 1988; 16: 5115–5122.
[11] Azhikina T, Veselovskaya S, Myasnikov V, Potapov V, Ermolayeva O, Sverdlov E. Proc Natl Acad Sci USA 1993; 90: 11460–11462.
[12] Azhikina TL, Shevchenko YO, Lebedev YB, Veselovskaya SV, Myasnikov VA, Potapov VK, Sverdlov ED. Dokl Acad Nauk USSR 1993; 330: 642–645.
[13] Maniatis T, Frisch EF, Sambrook J (editors). Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press, 1989.
[14] Obradovitch D, Borodin AM, Kopantzev EP, Wagner LL, Volik SV, Ermolaeva OD, Lebedev YB, Monastyrskaya GS, Sverdlov ED. Bioorganicheskaya Khimya 1994; 20: 919–931.
[15] Lukyanov S, Gurskaya N, Lukyanov K, Tarabykin V, Sverdlov E. Bioorganicheskaya Khimya 1994; 20: 701–704.
[16] Siebert P, Chenchik A, Kellogg D, Lukyanov K, Lukyanov S. Nucleic Acids Res 1995; 23: 1087–1088.
[17] Lindahl T. Nature 1993; 362: 709–715.
[18] Kinzler KW, Vogelstein B. Nucleic Acids Res 1989; 17: 3645–3653.
[19] Ludecke H-J, Senger G, Klaussen U, Horsthemke B. Nature 1989; 338: 348–350.
[20] Mueller PE, Wold B. Science 1989; 246: 780–786.
[21] Johnson DH. Genomics 1990; 6: 243–251.
[22] Saunders RDC, Glover DM, Ashburner M, Siden-Kiamos I, Louis C, Monastirioti M, Savakis C, Kafatos E. Nucleic Acids Res 1989; 17: 9027–9037.
[23] Froussard P. Nucleic Acids Res 1992; 20: 2900.
[24] Cui X, Li H, Goradia TM, Lange K, Kazazian Jr., Galas D, Arnheim N. Proc Natl Acad Sci USA 1989; 86: 9389–9393.
[25] Meyerhans A, Vartanian J-P, Wain-Hobson S. Nucleic Acids Res 1990; 18: 1687–1691.
[26] Brakenhoff RH, Schoenmakers JGG, Lubsen NH. Nucleic Acids Res 1991; 19: 1949.
[27] Lawlor DA, Dickel CD, Hauswirth WW, Parham P. Nature 1991; 349: 785–788.

Chart I: Structures of A·T and AP·T Base Pairs



Crick type base pair with thymine (Sowers *et al.*, 1986), polymerases are still able to discriminate between insertion of 2-aminopurine and adenine opposite thymine, but misinsertion frequencies for AP are much higher than for natural nucleotides on the order of 10–15% (Bessman *et al.*, 1974; Clayton *et al.*, 1979; Pless & Bessman, 1983). For natural nucleotides, differences in the relative rates of correct and incorrect insertion are very large, on the order of a factor of 10^3–10^5, in comparison with differences in rates within different sequence contexts, which may vary by as much as a factor of 10 or more (Mendelman *et al.*, 1989). Evaluating differences between correct and incorrect insertion for natural nucleotides within different sequence contexts would require measuring relatively small differences in the ratio of a very large number to a very small number. Variations in the misinsertion frequency of AP due to changes in the local DNA sequence are likely to be on the same order of magnitude as the misinsertion frequency, making AP a more sensitive probe for measuring the effects of local DNA sequence on insertion. The kinetics of insertion of dAPMP and dAMP opposite T and their relative insertion efficiencies were measured on synthetic primer/templates (Chart II) having identical sequence except for the base pair immediately 5′ to the incoming dNTP, so that the effects of nearest neighbor interactions could be evaluated in the absence of the other sequence context effects.

A second advantage of using AP as a probe is that it is much more fluorescent than natural nucleotides, and its fluorescence properties are sensitive to its environment. We have shown that changes in the fluorescence intensity of AP can be used to follow polymerase-catalyzed insertion. A comparison is made between two methods for measuring the kinetics of DNA polymerase catalyzed reactions. In the first method, changes in the fluorescence intensity of AP as it was incorporated into DNA were used to measure reaction kinetics on pre-steady-state and steady-state time scales. The second method uses a gel assay system (Boosalis *et al.*, 1987; Goodman *et al.*, 1993) to measure incorporation of both AP and A deoxyribonucleotides.

[1] Abbreviations: dAPTP, 2-aminopurine 2′-deoxyribonucleoside 5′-triphosphate; AP, 2-aminopurine; dNTP, 2′-deoxyribonucleoside 5′-triphosphate; KF⁻, a 3′→5′ exonuclease-deficient mutant of the large (Klenow) of *Escherichia coli* DNA polymerase I; KF⁺, Klenow fragment with the 3′→5′ exonuclease activity present; DTT, dithiothreitol; BSA, acetylated bovine serum albumin; DMT, dimethoxytrityl protecting group; CPG, controlled-pore glass.

Chart II: Sequence of Primer/Templates

primer templates for standing start reactions:

```
16mer:  5'- TCC CAG TCA CGA CGT Y
30mer:  3'- AGG GTC AGT GCT GCA XTZ GTA CGA GCT ACT
```

primer templates for running start reactions (universal primer, up):

```
15mer:  5'- TCC CAG TCA CGA CGT
30mer:  3'- AGG GTC AGT GCT GCA XTZ GTA CGA GCT ACT
```

primer (p)	Y	template (t)	X	Z
pG	G	tC	C	A
pC	C	tG	G	A
pA	A	tT	T	A
pT	T	tA_1	A	A
up		tA_2	A	G

EXPERIMENTAL PROCEDURES

Materials

KF⁻ (D424A) was a kind gift from Dr. Catherine M. Joyce (Derbyshire *et al.*, 1991). Klenow exo⁻ (D355A, E357A) was either purified (Derbyshire *et al.*, 1988; Joyce & Grindley, 1983) from an overproducing strain provided by Dr. Catherine M. Joyce or purchased from U.S. Biochemical Corp. KF⁻ solutions were diluted for storage in a buffer consisting of 20 mM Tris, pH 7.5, 0.5 mM dithiothreitol, 1 mg/mL BSA, and 50% glycerol. Concentrations of KF⁻ were determined by measuring the absorbance at 278 nm ($\epsilon = 6.32 \times 10^4$ M^{-1} cm^{-1} (Setlow *et al.*, 1972)), and concentrations of KF⁻ from US Biochemical Corp. were determined from specific activities and protein concentrations. Acetylated bovine serum albumin was purchased from U.S. Biochemical Corp. Restriction enzymes *Mbo*I, *Dde*I, and *Hin*fI were purchased from Pharmacia, and *Exo*III was purchased from Boehringer Mannheim. FPLC-purified dNTPs were purchased from Pharmacia LKB Biotechnology, Inc., and used without further purification. Oligonucleotides were synthesized on an Applied Biosystems Model 381A DNA synthesizer using standard 2-cyanoethyl phosphoramidite chemistry and purified by denaturing polyacrylamide gel electrophoresis. Nucleotide sequences are given in Chart II and referred to by their designations throughout the text. A DMT-protected 2-aminopurine was linked to a solid support (CPG) for preparation of oligos containing AP at the 3′-terminus (Eritja *et al.*, 1986). dAPTP was prepared as described previously (Bessman *et al.*, 1974; Clayton *et al.*, 1979). Concentrations of dAPTP were determined by measuring the absorbance at 304 nm ($\epsilon = 5800$ M^{-1} cm^{-1}).

Methods

Time-Resolved Measurements. Time-resolved experiments were performed with free dAPTP and with a synthetic primer/template of identical sequence to pT/tA_1 but containing AP at the 3′-primer terminus. Free dAPTP data was collected at 0.5 μM in 50 mM Tris-HCl, pH 7.4, and the AP primer/template data was collected in 25 mM HEPES and 50 mM NaCl. KF⁻–DNA complex measurements were made on a solution containing 1 μM KF⁻ (D355A, E357A), 0.5 μM pG/tC, 0.5 mM EDTA, 25 mM HEPES, pH 7.5, and 50 mM NaCl.

Time-resolved fluorescence measurements were performed utilizing a Coherent Antares Nd:YAG laser (Palo Alto, CA), frequency doubled and synchronously pumping a dual dye-jet laser (Coherent 702) using rhodamine 6G and a DODCI saturable absorber. Output pulses from this laser at 306 nm were utilized at 4 MHz and had a pulse width of approximately 1 ps. Time-resolved detection utilized time-correlated single-

photon counting with a ·m Hamamatsu (R2809U-01, Bridgewater, NJ) microch. ·l plate detector, high-frequency 50X amplifiers (Phillips Scientific 774, Mahwah, NJ), constant fraction discriminators (Tennelec 454, Oak Ridge, TN), time-to-amplitude converters (Tennelec 862), and pulse-height analysis analog-to-digital converters (Nucleus PCA-II, Oak Ridge, TN). The instrument response function was approximately 50–100 ps. The collimated fluorescence emission was passed through Glan-Thompson polarizers on automated mounts (ISS, Urbana, IL) and focused onto the entrance slits of a SPEX (Edison, NJ) 0.22-m emission monochromator. A half-wave plate in the excitation beam was utilized to rotate the excitation polarization to horizontal for the determination of the polarization bias ("*g*-factor") of the detection instrumentation.

Quantum Yield Determination. A SPEX Fluorolog 1681 fluorimeter was used to measure quantum yields. Emission spectra were collected for solutions of 4 μM primer/template or dAPTP in 25 mM HEPES and 50 mM NaCl, pH 7.5 between 330 and 460 nm at 0.5-nm intervals and integrated over 0.5 s. Buffer background signals were subtracted from spectra. Relative quantum yield values were calculated by dividing the integral of the primer/template fluorescence by that of the dAPTP fluorescence. Absolute quantum yield values were obtained by multiplying relative values by the reported quantum yield of 0.63 for dAPTP (Ward *et al.*, 1969).

Fluorescence Detection. Fluorescence detection from the stopped-flow apparatus consisted of a home-built single-photon detector consisting of the following: Hamamatsu R928 photomultiplier, 5X 300-MHz amplifier (Stanford Research SR445, Sunnyvale, CA), discriminator (Stanford Research SR400), and multichannel scaler (Tennelec Model MCS-II, Oak Ridge, TN), interfaced to an 80486 microcomputer. The detection system was activated from an external synch-out pulse from the Molecular Kinetic stepper-motor controlling unit, and data acquisition began at least 100 ms before sample mixing. Data was collected at time bases from 1 ms to 1 s in 8000 total channels. Fluorescence excitation utilized a 75-W xenon arc lamp (Oriel, Stratford, CT) coupled to a 0.25-m monochromator (Oriel 77250) with fiber optic output directed into the 75-μL-volume sample cell. Excitation wavelength was 311 nm, and emission was collected through a 360-nm cut-on filter (Hoya Optics type L36, Fremont, CA).

Fluorescence Determination of Steady-State Kinetic Constants Using Standing-Start Primer/Templates. Reactions were perfomed using rapid mixing techniques in a Molecular Kinetic (Pullman, WA) three-syringe stepper-motor-controlled stopped-flow apparatus. Rate of product formation was determined from the decrease in the steady-state dAPTP fluorescence. Apparent K_m and V_{max} values were obtained using the method described below under Calculation of Kinetic Constants. Stock solutions of dAPTP, a dilution buffer, and primer/template–KF^- were placed in separate syringes. Reactions were perfomed at 19 ± 2 °C by mixing an aliquot of dAPTP with dilution buffer and then mixing 300 μL of the diluted dAPTP solution with 150 μL of primer/template–enzyme stock solution. Volumes of the dAPTP and dilution buffer solutions were varied between 0 and 300 μL to give desired dAPTP concentrations, which varied as per template and are listed below. All reactions contained final concentrations of 50 mM Tris-HCl, pH 7.4, 5 mM $MgCl_2$, 0.3 mM DTT, 0.03 mg/mL BSA, 5 nM KF^- (D355A, E357A), and 70 nM standing-start primer/template (Chart II). Template tA_2: Low- (0.75 μM) and high-concentration dAPTP (2.25 μM) stock solutions were used to ramp the concentration from 0.1 to 0.5 and from 0.6 to 1.5 μM, respectively. Template tC: A 1.5 μ. dAPTP stock solution was used to vary the concentration from 0.2 to 1.0 μM. Templates tA_1, tT, and tG: A 0.75 μM dAPTP stock solution was used to ramp concentrations from 0.1 to 0.5 μM, and a 3.75 μM dAPTP stock solution was used for 0.75 to 2.5 μM dAPTP concentrations. For these last three templates multiple runs (2–5) were summed to increase signal to noise.

Saturated Steady-State Kinetics. A more fully saturated steady-state experiment was run using the pC/tG primer/template. Four dAPTP stocks were used in this experiment of 2.0, 8.0, 16, and 24 μM to vary dAPTP concentration from 0.5 to 12 μM. Reactions were otherwise identical with the steady-state experiments above except that the pC/tG concentration was 0.5 μM. At the higher dAPTP concentrations multiple (2–5) runs were summed to increase signal to noise.

Fluorescence Determination of Pre-Steady-State Kinetic Constants. Reactions were performed in the stopped-flow apparatus at 19 ± 2 °C. Primer/template–KF^- (D355A, E357A) solutions were prepared which contained 0.1 mg/mL BSA and 1 mM DTT. Reactions were initiated by mixing equal volumes (150 or 180 μL) of primer/template–KF^- solution with dAPTP/dilution buffer solution. All reactions contained 50 mM Tris-HCl, pH 7.4, and 8 mM $MgCl_2$. Final concentrations of dAPTP in reactions ranged from 0.6 to 1.8 μM for pC/tG and 0.6 to 3.6 μM for pT/tA_1. Final concentrations of primer/template and KF^- were 300 and 100 nM. Pre-steady-state kinetics were obtained at high dAPTP concentrations for primer/templates pC/tG and pT/tA_1 using 0.2 μM Klenow fragment, 0.5 μM DNA, and 10 μM dAPTP. Reactions were initiated as above by mixing 120 μl of 0.4 μM Klenow fragment and 1.0 μM DNA mixed with 120 μl of 20 μM dAPTP. Multiple experiments were summed (approximately 30 runs) to obtain acceptable signal to noise.

Single-Turnover Pre-Steady-State Experiment. Reaction conditions were as for pre-steady-state experiments unless stated otherwise. Final concentrations were 1.0 μM and 0.5 μM for KF^- (D355A, E357A) and pT/tA_1, respectively. Four dAPTP stocks of 2.0, 8.0, 16, and 24 μM were used to ramp dAPTP concentrations from 0.5 to 12 μM. All velocities above 4 μM dAPTP were summed results from multiple experiments.

Calculation of Kinetic Constants. All fluorescence reactions were normalized to total change in AP fluorescence because the amount of AP fluorescence quench seen varied between the four different primer/templates. Initial velocities were obtained directly by calculating the slope of the initial part of the reactions. For steady-state reactions, k_{cat} and K_m values were calculated from plots of observed initial rates versus dAPTP concentration. For pre-steady-state reactions, kinetic constants were calculated by assuming the pathway below, which is described by the rectangular hyperbola in eq 1:

$$KF^-\cdot DNA + dAPTP \overset{K_m}{\rightleftharpoons} KF^-\cdot DNA\cdot AP \overset{k_{cat}}{\rightarrow} KF^-\cdot DNA_{+1}\cdot PP_i$$

$$k_{obs} = \frac{k_{cat}[dAPTP]}{K_m + [dAPTP]} \approx \frac{k_{cat}}{K_m}[dAPTP] \qquad (1)$$

At high concentrations of dAPTP in comparison to DNA, the change in the large dAPTP signal was small and the resulting noise in data points was large. At lower concentrations of dAPTP where individual data points are more reproducible, plots of k_{obs} versus [dAPTP] showed a small amount of curvature. Because of the signal-to-noise problems, separate values of k_{cat} and K_m were not extracted. Instead plots of k_{obs}

versus [dAPTP] were fit to the linear approximation in eq 1. [illegible] by using pre-steady-state kinetic measurements were similar to those obtained from steady-state measurements.

Preparation of DNA Trap. Calf thymus trap DNA was prepared as described (Joyce, 1989) except that the DNA was sheared before enzymatic digestion by passing it through a narrow-bore needle several times and that the restriction enzyme *Mbo*I was substituted for *Sau*3AI.

Single-Hit Kinetic Measurements in the Presence of Trap DNA. Primer 5′-end labeling using 0.5 equiv of [γ-^{32}P]ATP, annealing, electrophoresis, and data analysis were done as described by Boosalis *et al.* (1987) except that a Molecular Dynamics PhosphorImager was used to determine the intensities of gel bands rather than autoradiography and densitometry. Nucleotide solutions for kinetics were prepared by mixing a 10× Klenow reaction buffer consisting of 0.50 M Tris, pH 7.4 at 20 °C, 60 mM $MgCl_2$, and 5 mM dithiothreitol with 10× solutions of dNTPs and dH_2O to give 1× concentrations when finally diluted in kinetics reactions. KF$^-$ (D424A) was diluted 1:10 in a solution of 0.2 mg/mL BSA and 1 mM dithiothreitol. Kinetics reactions were initiated by mixing 4 μL of a dNTP/trap DNA solution consisting of equal volumes of dNTP solution and a solution of 4 mg/mL trap DNA with 4 μL of primer/template–Klenow solution consisting of equal volumes of labeled DNA and diluted KF$^-$ solution. After 15 s at 20 °C, reactions were quenched with 24 μL of 20 mM EDTA in 95% formamide. Final concentrations of KF$^-$ (D424A), primer/template, and trap DNA were 23 nM, 50 nM, and 1 mg/mL, respectively. Final concentrations of running-start bases were 40 μM for dTTP, 20 μM for dCTP, and 5.5 μM for dGTP. Concentrations of dATP were varied between 0.085 and 19 μM, and concentrations of dAPTP were varied between 0.11 and 57 μM.

Control reactions were performed with up (universal primer)/tA_2 by incubating the reactions for 5, 15, 30, and 45 s to ensure that a 15-s reaction time was long enough for the polymerase reaction to be completed. Control reactions were performed to determine the effectiveness of the DNA trap by preincubating the DNA primer/template and trap DNA before adding polymerase and dNTP solutions. The percentages of primer extension observed when the primer/templates were preincubated with trap DNA were not greater than 2% of the percentages of primer extension observed in kinetics reactions.

Reactions with the KF$^-$ double mutant (D355A, E357A) contained 12 nM KF$^-$, 50 nM primer/template, and 1 mg/mL trap DNA. Polymerase was diluted 1:10 in a solution that contained final concentrations of 0.18 mg/mL BSA, 0.7 mM DTT, and 2× Klenow reaction buffer. Nucleotide/trap DNA solutions and KF$^-$/DNA solutions were made as above, and reactions were performed as described above. Final concentrations of the running-start bases were 6.7 μM for dTTP, 6.9 μM for dCTP, and 6.8 μM for dGTP. Concentrations of dATP ranged from 0.085 to 18.7 μM, and concentrations of dAPTP ranged from 0.22 to 57 μM.

Measurement of Polymerase–DNA Dissociation Rates. Nucleotide and primer/template–polymerase solutions were prepared as described for the kinetics measurements in the presence of trap DNA. Reactions were initiated by mixing 4 μL of the running-start dNTP/trap DNA solution with 4 μL of the primer/template–Klenow solutions at 20 °C. After delay times between 3 and 150 s, a solution of a saturating concentration of dATP in 1× Klenow reaction buffer was added to the reaction mixture, and the reaction was quenched after 15 s with 30 μL of 20 mM EDTA in 95% formamide. For a 0-s delay time reaction, a dNTP/trap DNA solution was prepared which contained both the running-start dNTP and dATP, and 4 μL of this solution was added to 4 μL of primer/template–enzyme solution before the reaction was quenched at 15 s. Final concentrations of KF$^-$ (D424A), primer/template, and trap DNA were 23 nM, 50 nM, and 1 mg/mL, respectively. Measurements were made twice for each primer/template. Nucleotide concentrations were as follows for the different primer/templates: 40 μM dTTP and 50 or 51 μM dATP for up/tA_1 and up/tA_2, 20 μM dCTP and 26 or 51 μM dATP for up/tG, and 5.6 μM dGTP and 26 or 51 μM dATP for up/tC. For reactions with the KF$^-$ double mutant (D355A, E357A), final concentrations of KF$^-$, primer/template, and trap DNA were 12 nM, 50 nM, and 1 mg/mL, respectively. Each reaction mixture contained 51 μM dATP and the following concentrations of running-start nucleotide: 6.7 μM dTTP for up/tA_1, 6.9 μM dCTP for up/tG, and 6.8 μM dGTP for up/tC.

Values for the dissociation rate constants, k_{off}, were determined by fitting the fraction of primers extended from the site before the target site (site $i-1$) to the target site (site i) as a function of time to a first-order exponential decay (see Figure 4). Control reactions in which primer/templates were preincubated with trap DNA before dNTPs and polymerase were added were performed with delay times between 10 and 120 s to ensure that the trap was effective throughout the entire time course of the reactions. The percentages of primers extended in the control reactions at the longest delay times used in the reactions were not greater than 2% of the percentages of primers extended in k_{off} reactions.

Measurements of k_{off} in the presence of GTP which is not incorporated were done as above, using 25 nM KF$^-$ (D355A, E357A), 50 nM up/tA_1, 1 mg/mL BSA, 49 μM dTTP, and 51 μM dATP. GTP was added to the dTTP/trap DNA solutions to give final concentrations of 0, 60, and 250 μM GTP. KF$^-$ was diluted 1:10 in a solution containing final concentrations of 0.18 mg/mL BSA and 0.9 mM DTT before it was added to primer/template.

Direct Competition. Primer 5′-end labeling, annealing, electrophoresis, and data analysis were done as described above for insertion kinetics. Nucleotide/trap DNA solutions and KF$^-$ (D355A, E357A)/primer/template solutions were prepared as for the separate kinetics measurements, and the general procedure was the same as that used for separate kinetics. Final concentrations of KF$^-$, up/tC, trap DNA, and dGTP were 12 nM, 50 nM, 1 mg/mL, and 6.8 μM, respectively. Each reaction mixture contained 0.17 μM dATP and concentrations of dAPTP ranging from 0.11 to 14.5 μM. Reactions were quenched after 15 s. Primers ending in either A or AP can be separated by running on a 16% polyacrylamide gel at room temperature; see Figure 5.

The misinsertion efficiency of AP, $f_{ins} = (k_{cat}/K_m)_{AP}/(k_{cat}/K_m)_A$, can be calculated from the fraction of primers extended by A in the presence of competing AP as derived below. The fraction of primers extended by dAMP in the presence of dAPMP is given by the relative rates of insertion of A and AP, as in eq 2,

$$\frac{I_A}{I_A + I_{AP}} = \frac{(k_{cat}/K_m)_A[A]}{(k_{cat}/K_m)_A[A] + (k_{cat}/K_m)_{AP}[AP]} \quad (2)$$

where $I_{A\ and}\ I_{AP}$ are the band intensities for primers ending in A and AP, respectively, and [A] and [AP] are the concentrations of dATP and dAPTP, respectively.

Misinsertion efficiencies were calculated by rearranging eq 2 to eq 3 and fitting the observed band intensities at each

dAPTP concentration to eq 3.

$$\frac{I_A}{I_A + I_{AP}} = \frac{[A]}{[A] + \frac{(k_{cat}/K_m)_{AP}}{(k_{cat}/K_m)_A}[AP]} \quad (3)$$

The misinsertion efficiency calculated from direct competition using eq 3 can be compared to the values of k_{cat}/K_m obtained by measuring the incorporation of dAMP as a function of [dATP] and of dAPMP as a function of [dAPTP] in separate experiments, i.e., in the absence of direct competition.

RESULTS

To investigate the effects of nearest neighbor interactions on nucleotide insertion kinetics and fidelity, we measured the misinsertion of dAPMP and the correct insertion of dAMP opposite a template T in synthetic oligonucleotide primer/templates of identical sequence except for the nucleotide at the 3′-primer terminus (5′-nearest neighbor). Kinetic studies of insertion were done by two methods. The first method monitored changes in the fluorescence of AP during polymerase-catalyzed incorporation of dAPMP at millisecond intervals. The second method used a gel assay to measure the kinetics for insertion of dAPMP or dAMP opposite T (Boosalis *et al.*, 1987; Goodman *et al.*, 1993), from which the AP misinsertion efficiency can be deduced (Fersht, 1985). Alternatively, the gel assay can be used to measure the AP misinsertion efficiency directly by allowing the two substrates to compete for insertion into DNA.

Insertion of dAPMP by KF⁻ was measured next to the four common 5′-nearest neighbor nucleotides. Oligonucleotide sequences are shown in Chart II. Insertion of dAPMP opposite T in an oligonucleotide primer/template that differs in the identity of the base 5′ to the template T site (template tA_2) was also examined to measure the effect of cross stacking on insertion. The majority of the fluorescence measurements were done using standing-start primer/templates (Chart II) where the oligonucleotide primer ends at the site immediately before the template target T site. In the gel assay, running-start primer/templates (Chart II) were used where insertion of one nucleotide is required to extend the primer to the target T site. Mutants of the large fragment of DNA polymerase I which were deficient in their 3′→5′ exonuclease activity, KF⁻, were used in this study so that the insertion kinetics were not complicated by exonucleolytic removal of inserted nucleotides. Two different exonuclease deficient mutants, one having a single amino acid substitution (D424A) and the other having two amino acid substitutions (E355A, D357A), were examined.

Fluorescence Properties of 2-Aminopurine. The fluorescence emission of AP is sensitive to its environment (Ward *et al.*, 1969). When AP is present as a deoxyribonucleoside triphosphate, its steady-state emission is about 25–125 times greater, depending on surrounding sequence, than when present at a primer 3′-terminus. When AP is present at the 3′-terminus of double-stranded DNA, its emission is quenched relative to single-stranded DNA (Figure 1a). The excitation maximum for AP is at 310 nm, and the emission maximum is at 365 nm, for both the free nucleotide and the nucleotide incorporated into DNA.

The fluorescence emission intensity is sensitive to the nearest neighbor base-stacking partner for AP. The intensities decrease as the base-stacking partner for AP is changed in the order T ≈ A > C > G. Quantum yields for AP within these primer/templates relative to free dAPTP are 3.7, 3.5, 2.4,

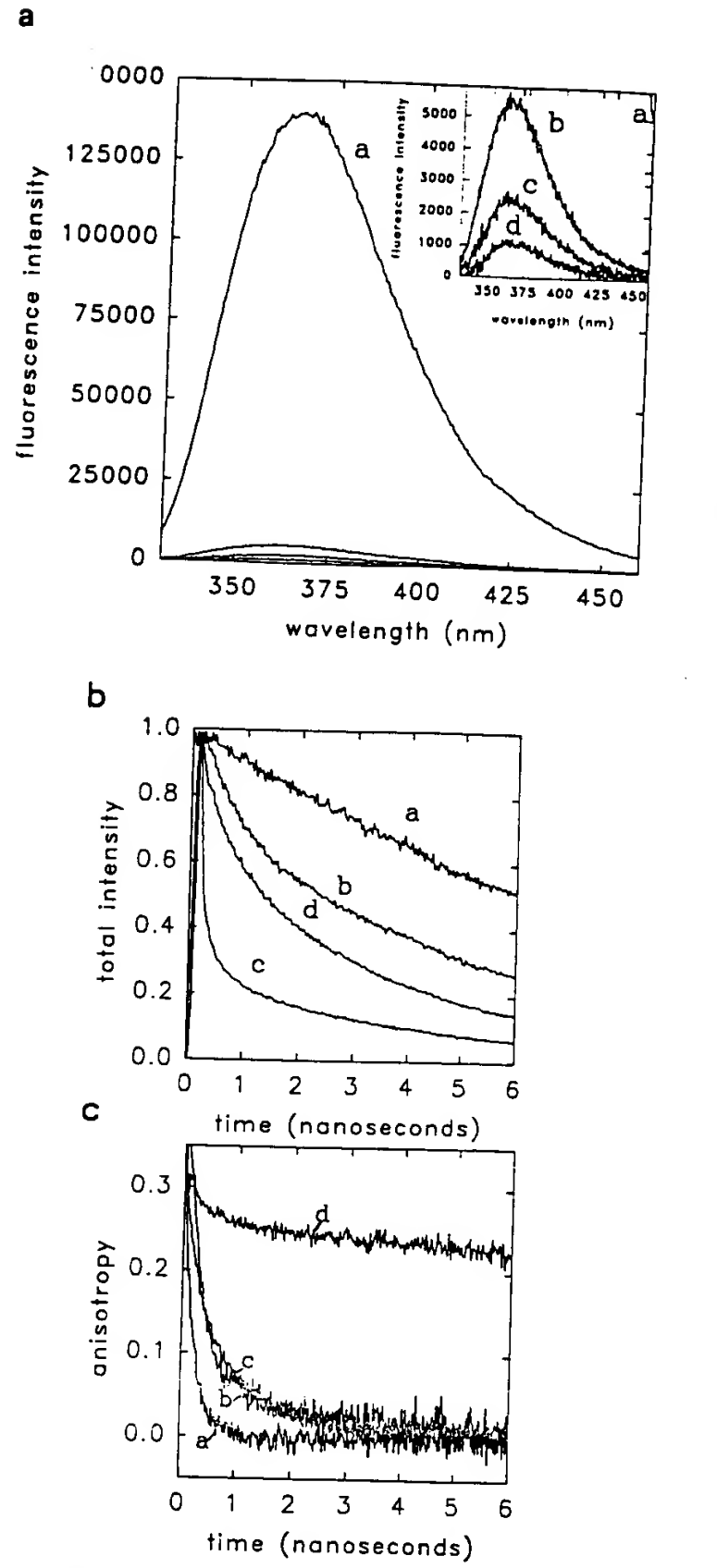


FIGURE 1: Steady-state fluorescence emission spectra and time-resolved total intensity and anisotropy decays for AP in different environments. Conditions are described in Methods. (Panel a) Steady-state fluorescence emission spectra for dAPTP (a), a synthetic single-stranded oligo of identical sequence to pG but containing AP at the 3′-terminus (b), KF⁻(D355A, E357A) bound to a duplex made up of the AP primer in spectrum b annealed to template tC (c), and the synthetic AP-containing duplex free in solution (d). The inset shows the less intense spectra on an expanded scale. (Panel b) Time-resolved total intensity decays for dAPTP (a), a synthetic 17mer oligo of identical sequence to pC with the addition of AP at the 3′-terminus (b), a duplex consisting of the primer in spectrum b annealed to template tG (c), and a duplex of the same sequence as pG/tC containing AP at the 3′-terminus bound to KF⁻ (d). (Panel c) Time-resolved anisotropy decays for the AP species in panel b.

and 0.8% for tA_1, tT, tG, and tC templates, respectively. Using the quantum yield of 63% reported for dAPTP (Ward *et al.*, 1969) gives absolute quantum yields for AP within these primer/templates of 2.3, 2.2, 1.5, and 0.5% for templates tA_1, tT, tG, and tC, respectively. Time-resolved fluorescence emission and anisotropy decays for AP are sensitive to the environment. The fluorescence intensity of free dAPTP in solution decays as a single exponential with a lifetime of 8.57 ns compared to 10.43 ns previously reported for the nucleoside

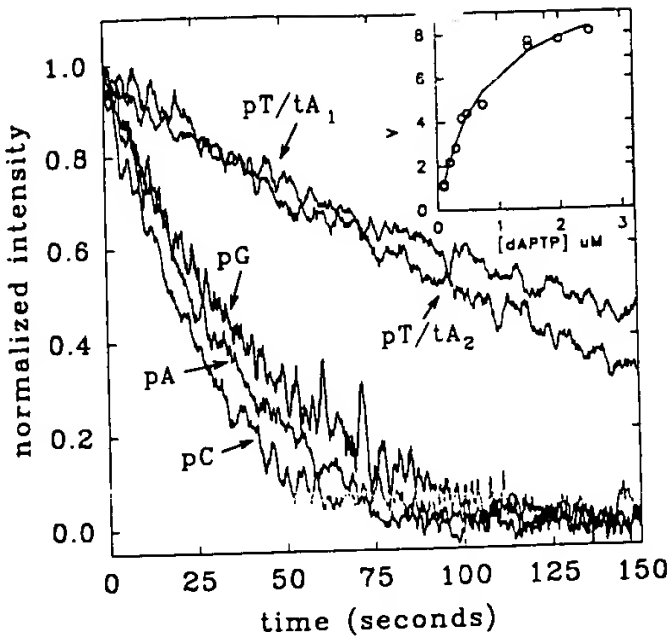


FIGURE 2: Time courses for steady-state incorporation of dAPMP into each of the five primer/templates. Reactions were performed at 20 °C using 5 nM KF⁻ (D355A, E357A), 70 nM primer/templates, and 0.2 μM dAPTP as described in Methods. The inset shows a plot of observed reaction rate versus concentration of dAPTP for reactions containing 5 nM KF⁻ (D355A, E357A) and 70 nM pC/tG as described in Methods.

(Guest *et al.*, 1991b). The triphosphate group most likely quenches the AP fluorescence to some extent. When AP is present at the 3′-primer terminus in duplex DNA, there is a large decrease in the fluorescence lifetime (Figure 1b). Four lifetime components are required to fit the data, similar to previous reports for AP at the center of duplex DNA (Guest *et al.*, 1991b; Nordlund *et al.*, 1989). The fluorescence lifetime increases for AP at the 3′-primer terminus of duplex DNA when AP is bound to KF⁻ (D355A, E357A), which corresponds with the ≈2.5-fold enhancement of steady-state fluorescence for the KF⁻–DNA complex over free DNA (Figure 1a,b).

A single rotational correlation time (0.163 ns) was obtained from free dAPTP. When located at the 3′-primer terminus in DNA duplexes, AP exhibited long (~3–6 ns) and short (0.15–1 ns) rotational correlational times which compare well with a previous study (Guest *et al.*, 1991b) (Figure 1c). The shorter correlation times are most likely due to AP base motion, while the longer times can be ascribed to some component of the overall motion of the 17/30mer primer/templates. There is a dramatic increase in the long rotational correlation time for double-stranded DNA containing AP at the 3′-primer terminus when bound to KF⁻ (D355A, E357A). This long rotational correlation time (~42 ns) probably reflects the overall motion of the large (69 kD) KF⁻–DNA complex. A shorter rotational correlation time is also present for the complex, and it is similar in magnitude to the one found for double-stranded DNA alone and probably reflects AP base motion within the complex.

Steady-State and Pre-Steady-State Incorporation of dAPMP Measured by Fluorescence Spectroscopy. The kinetics of insertion of dAPMP by KF⁻ can be measured by following the decrease in the steady-state fluorescence emission of AP during the time course of a reaction. Data for reactions on each primer/template were normalized to the total change in AP fluorescence. The difference in insertion rates on the five templates can be seen clearly in Figure 2. Changing the downstream neighbor on the template from A (Figure 2, tA_1) to G (Figure 2, tA_2) does not affect the insertion efficiency significantly. Rates were calculated by linear fits to data from reaction time courses at early times. Kinetic constants were calculated from plots of the observed rate of insertion versus dAPTP concentration. A typical plot is shown in the inset to Figure 2, and kinetic constants for insertion of dAPMP are given in Table I. The relative rates of insertion of dAPMP, given by the ratios of [dAPTP]k_{cat}/K_m on different primer/templates, are slowest next to T. Changes in the relative insertion efficiencies of dAPMP next to different 5′-nearest neighbors result primarily from changes in k_{cat}. Values of apparent K_m range from 0.78 to 1.1 μM, while k_{cat} values range from 0.41 s^{-1} next to T to 2 s^{-1} next to G. Similar results were obtained by using higher [dAPTP] and DNA/enzyme ratios. Values of k_{cat} and K_m for insertion on pC/tG were 0.95 s^{-1} and 1.3 μM, respectively, in reactions containing 5 nM KF⁻ (D355A, E357A), 500 nM DNA, and 0.5–12 μM dAPTP.

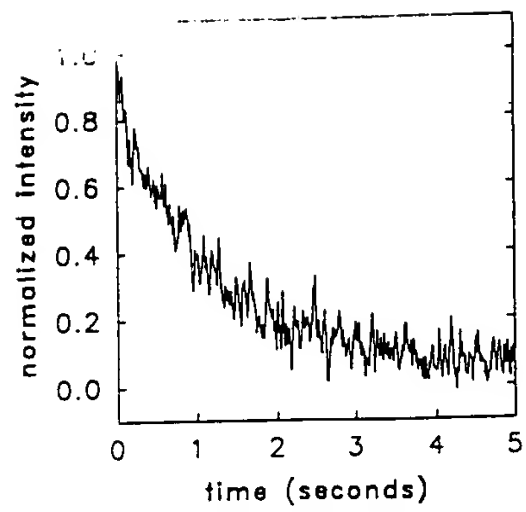


FIGURE 3: Pre-steady-state kinetics of insertion of dAPMP into pC/tG by KF⁻ (D355A, E357A) at 20 °C as described inMethods. Concentrations of KF⁻, DNA, and dAPTP were 200 nM, 500 nM, and 10 μM, respectively.

Table I: Steady-State Kinetic Constants for Insertion of dAPMP opposite a Template T Determined by Fluorescence Spectroscopy[a]

primer/template	5′-nearest neighbor	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (μM^{-1} s^{-1})	slope[b] (μM^{-1} s^{-1})
pG/tC	G	1.1 ± 0.40	2.0 ± 0.49	1.8	1.3
pC/tG	C	0.78 ± 0.09	1.5 ± 0.08	2.0	1.4
pA/tT	A	0.81 ± 0.11	1.3 ± 0.08	1.5	1.2
pT/tA_1	T	0.90 ± 0.32	0.41 ± 0.06	0.46	0.59
pT/tA_2	T	0.82 ± 0.31	0.48 ± 0.20	0.58	0.45

[a] Reactions were performed at 20 °C as described in Methods using 5 nM Klenow exo⁻ (D335A, E357A) and 70 nM primer/template. Refer to Chart II for oligonucleotide sequences. [b] Values are listed for the slope of the linear region of [dAPTP] vs rate plots.

Kinetics of insertion of dAPMP were measured for two primer/templates (pC/tG and pT/tA_1) under pre-steady-state conditions (see Experimental Procedures) where KF⁻ and primer/template were preincubated to form an enzyme–DNA complex before dAPTP was added (Figure 3). The relative efficiency of insertion given by the slope of the linear region of plots of rate versus [dAPTP] was greater next to nearest neighbor C (k_{cat}/K_m = 0.60 s^{-1}) than next to nearest neighbor T (k_{cat}/K_m = 0.27 s^{-1}). These rates are on the same order of magnitude as in steady-state experiments and show the same trend (i.e., insertion next to C is more efficient than insertion next to T).

The change in the steady-state fluorescence of 2-aminopurine during an enzyme-catalyzed reaction can be attributed to incorporation of free 2-aminopurine into a DNA primer/template and not to alternative modes of quenching, for the following reasons. (1) The decrease in AP fluorescence follows the stoichiometry of the reactions. For example, if a reaction mixture contained 1.2 μM dAPTP and 300 nM DNA, the observed decrease in the AP fluorescence was consistent with the incorporation of 1/4 of the dAPTP into DNA. (2) The

Table II: Kinetic Constants and Misinsertion Frequencies for Insertion of dAMP or dAPMP opposite a Template T by Klenow exo⁻ (D424A) Determined Using the Gel Assay with a DNA Trap[a]

primer/ template	5′-nearest neighbor	dNMP	K_m (μM)	$(I_2/I_1)_{max}$	k_{cat} (s^{-1})	k_{cat}/K_m (μM^{-1} s^{-1})	f_{ins}
up/tC	G	A	0.97 ± 0.44	22 ± 3	5.6 ± 0.7	6.8 ± 2.4	1
		AP	2.7 ± 0.8	19 ± 2	5.0 ± 0.6	2.0 ± 0.4	0.31 ± 0.06
up/tG	C	A	0.93 ± 0.37	21 ± 6	3.8 ± 1.1	4.3 ± 0.5	1
		AP	4.6 ± 2.6	22 ± 9	4.0 ± 1.7	1.0 ± 0.2	0.23 ± 0.02
up/tA_1	T	A	0.48 ± 0.13	56 ± 11	2.1 ± 0.4	4.5 ± 0.4	1
		AP	2.7 ± 0.1	46 ± 5	1.7 ± 0.2	0.63 ± 0.05	0.15 ± 0.03
up/tA_2	T	A	0.34 ± 0	38 ± 5	0.83 ± 0.10	2.4 ± 0.3	1
		AP	2.5 ± 0.4	32 ± 5	0.69 ± 0.10	0.28 ± 0	0.12 ± 0.02

[a] Reactions were performed at 20 °C as described in Methods using 23 nM Klenow exo⁻ (D424A) and 50 nM running-start primer/templates. Results are the average of two separate experiments. Refer to Chart II for oligonucleotide sequences.

Table III: Kinetic Constants and Misinsertion Frequencies for Insertion of dAMP or dAPMP opposite a Template T by Klenow exo⁻ (D355A, E357A) Determined Using the Gel Assay with a DNA Trap[a]

primer/ template	5′-nearest neighbor	dNMP	K_m (μM)	$(I_2/I_1)_{max}$	k_{cat} (s^{-1})	k_{cat}/K_m (μM^{-1} s^{-1})	f_{ins}
up/tC[b]	G	A	0.41 ± 0.06	19 ± 0	4.6 ± 0	11 ± 2	1
		AP	1.8 ± 0.1	16 ± 1	4.1 ± 0.2	2.3 ± 0.02	0.20 ± 0.02
		A vs AP[c]					0.23 ± 0.01
up/tG	C	A	0.31	14	1.6	5.0	1
		AP	1.9	16	1.8	0.92	0.18
up/tA_1	T	A	0.20	59	0.64	3.2	1
		AP	0.95	45	0.50	0.52	0.16

[a] Reactions were performed at 20 °C as described in Methods using 12 nM Klenow exo⁻ (D355A, E357A) and 50 nM running-start primer/templates. Refer to Chart II for oligonucleotide sequences. [b] Results are the average of two separate experiments. [c] Results from direct competition reaction under the same conditions as separate kinetic reactions, as described in Methods.

decrease in AP fluorescence follows the stoichiometry of the reaction regardless of whether the polymerase is present in stoichiometric amounts with the DNA or the DNA is in a large excess, indicating that there is no nonspecific binding of the AP to the polymerase causing the observed quenching. (3) In a reaction containing 100 nM dAPTP, 80 nM up/tC, and 5 nM KF⁻ where dAPMP was *not* incorporated opposite a C in template tC (data not shown), no decrease in AP fluorescence was observed, which further indicates that there is no nonspecific binding of AP to enzyme–DNA or bleaching of AP on the time scale of the reaction that leads to a decrease in AP fluorescence. (4) In reactions where dATP was allowed to compete with dAPTP for insertion (data not shown) both the observed rate of dAPMP incorporation and the total incorporation of dAPMP decreased as predicted from the gel competition assay (see Table III). (5) Time-resolved rotational anisotropies and total intensity decays measured on completed reactions containing an excess of DNA over enzyme are consistent with those observed for AP at the 3′-primer terminus of synthetic primer/templates.

The differences in the intensities of AP at the 3′-terminus of a synthetic duplex free in solution or bound to KF⁻ are small compared to the intensity of dAPTP. AP fluorescence increases by 2.5-fold when a synthetic primer/template containing AP at the 3′-primer terminus is bound to KF⁻ (Figure 1a), compared with a 25–125-fold decrease in fluorescence between free dAPTP and AP in duplex DNA. Thus, the observed quench in AP fluorescence does not arise as a result of release of the DNA product containing AP from the enzyme to an aqueous environment. We cannot rule out the possibility that the fluorescence of AP is quenched when it binds to the enzyme–DNA complex in a conformation where it is base paired with the template, stacked with the primer terminus, and poised for phosphodiester bond formation. It seems likely that the AP insertion rates may reflect a slower step than phosphodiester bond formation, e.g., a conformational change that orients the substrates in position to react (Kuchta *et al.*, 1987).

Biphasic reaction kinetics, which would indicate that the first turnover of enzyme–DNA substrate to enzyme–DNA product was faster than subsequent rounds of synthesis, were not observed for any of the DNA sequences. The rate constants for polymerization (k_{cat} values) are similar for both the steady-state and the pre-steady-state experiments, indicating that both types of experiment are measuring the same kinetic step. In a single-turnover experiment containing 1 μM KF⁻ (D355A, E357A) and 0.5 μM DNA (pT/tA_1), a value of k_{cat}/K_m = 0.63 s^{-1} was measured, which is similar to that obtained in the steady-state experiments (Table I). These results indicate that the observed rate is less than or equal to the rate of release of DNA product from polymerase.

Gel Assay Run under "Single Completed Hit" Conditions. Kinetics of insertion for both dAPMP and dAMP next to different base-stacking partners were measured using a gel assay (Boosalis *et al.*, 1987; Goodman *et al.*, 1993). When reactions are performed with a large excess of DNA primer/template in comparison with polymerase so that fewer than 20% of the labeled primers are extended, gel bands arise from a single encounter of a primer/template with a polymerase (single-hit conditions) (Goodman *et al.*, 1993). Under these single-hit conditions, fewer than 2% of primer/templates are extended more than once. Using the primer/template in large excess over the polymerase also ensures that observed band intensities contain an insignificant contribution from quenching of the reactions before the polymerase has completed synthesis on a given primer/template, thus stopping a polymerase prematurely (completed-hit conditions).

When polymerase reactions are performed under "single completed hit" conditions, the polymerase either inserts a nucleotide opposite the target site, measured by the intensity of the gel band at the target, or dissociates from the template prior to reaching the target, measured by the intensity of the

gel band at the template site prior to the target (Goodman *et al.*, 1993). The ratio of the band intensities (I_i/I_{i-1}) at the target site (site i) and the site immediately before the target site (site $i-1$) is a measure of the relative rates of extension of a primer from site $i-1$ to site i and dissociation at site $i-1$.[2] Another way to achieve single completed hit conditions is to run the reactions in the presence of a large excess of unlabeled trap DNA so that when the polymerase dissociates from the labeled DNA primer/template, it interacts with the large excess of unlabeled trap DNA rather than reassociating with labeled primer/template DNA. A plot of the observed ratio of band intensities, $(I_i/I_{i-1})_{obs}$, versus [dNTP] follows Michaelis–Menten kinetics and yields values for $(I_i/I_{i-1})_{max}$ and apparent K_m (Boosalis *et al.*, 1987):

$$(I_i/I_{i-1})_{obs} = \frac{(I_i/I_{i-1})_{max}[\text{dNTP}]}{[\text{dNTP}] + K_m} \quad (4)$$

The quantity $(I_i/I_{i-1})_{max}$ measures the ratio of the forward polymerization rate constant, k_{cat}, to the dissociation rate constant, k_{off} (at site $i-1$) (Goodman *et al.*, 1993). Thus, by measuring polymerase–DNA dissociation rates at site $i-1$, as described in the following section, we can calculate the slowest step, k_{cat}, along the pathway to extend a primer from site $i-1$ to i using eq 5.

$$k_{cat} = (I_i/I_{i-1})_{max}k_{off} \quad (5)$$

Mechanistic possibilities for this step include enzyme conformational changes (Kuchta *et al.*, 1987), translocation along the primer/template, and phosphodiester bond formation (Kuchta *et al.*, 1988). Product release rates cannot contribute to k_{cat} when reactions are carried out under single completed hit conditions.

Results for kinetics of insertion measured in the presence of a DNA trap are similar for both exonuclease-deficient mutants of Klenow (Table II for the single mutant, D424A, and Table III for the double mutant, D355A, E357A). Rate constants for dAMP insertion (k_{cat}), calculated by using eqs 4 and 5, are similar to those for insertion of dAPMP on the same primer/template for each polymerase. Values of K_m for insertion of dAPMP are about 4–8-fold larger than K_m values for insertion of dAMP, while k_{cat} values differ by only about 5–30%. Thus, misinsertion of dAPMP depends more on differences in K_m than on differences in k_{cat}. In agreement with the fluorescence data for insertion of dAPMP (Figure 2), rates of insertion of both dAPMP and dAMP are faster following a primer terminus G or C than a primer terminus T.

Polymerase Dissociation Rates from Primer/Templates. In order to calculate k_{cat} (eq 5) and make comparisons between k_{cat} values for insertion of dNTPs on primer/templates of different sequences, it is necessary to measure k_{off} since this rate constant changes with different DNA sequences. For these reactions, a solution of KF$^-$ was preincubated with a 5′-^{32}P-labeled primer annealed to unlabeled template before a solution of the unlabeled running-start dNTP and an excess of trap DNA were added. After a delay time between 0 and 150 s, a solution of a saturating concentration (51 μM) of dATP was added. During the delay time, some of the enzyme remained bound to the primer/template while some of the enzyme dissociated and interacted with the excess trap DNA. The polymerase that remained bound to primer/template was extended to the target T site in the presence of dATP. A gel showing the results of a typical k_{off} determination appears in Figure 4a. The fraction of polymerase remaining bound to the primer/template at any given delay time is equal to $I_i/(I_i + I_{i-1})$. A plot of $I_i/(I_i + I_{i-1})$ versus the delay time decays as a first-order exponential with rate constant k_{off} (Figure 4c) (Goodman *et al.*, 1993).

Dissociation rates were measured for both KF$^-$ mutants. The rate of dissociation (measured in two separate experiments for each primer/template) of KF$^-$ (D424A) is slightly faster when the primer ends in G (0.26 ± 0.02 s^{-1}) than when it ends in C (0.18 ± 0.01 s^{-1}), and about 7 times faster when the primer ends in G than in T (0.037 ± 0.006 s^{-1}), in reactions containing 5.6, 20, and 40 μM running-start dGTP, dCTP, and dTTP, respectively. Changing the base downstream of the target site from A (up/tA$_1$) to G (up/tA$_2$) reduces k_{off} by a factor of 1.7 to 0.022 ± 0.001 s^{-1}. Dissociation rates for the double mutant (D355A, E357A) show the same trend. Dissociation is faster following a primer terminus G (0.24 ± 0.02 s^{-1}) than following C (0.11 s^{-1}), and about 20 times faster following G than following T (0.011 s^{-1}), in reactions where running-start dNTP concentrations were 6.8, 6.9, and 6.7 μM for dGTP, dCTP, and dTTP, respectively. Large differences in dissociation rates also occur at different sites on the same primer/template. Dissociation of KF$^-$ (D424A) after insertion of dTMP (Figure 4a) on template tA$_1$ occurs at a rate of 0.037 ± 0.006 s^{-1}, while dissociation one base downstream after insertion of dAMP (Figure 4b,c) occurs at a rate we estimate to be at least 0.6 s^{-1}, or at least 16-fold faster.

Insertion Fidelity. Misinsertion efficiency (f_{ins}), given by the relative rates of insertion of incorrect (w) and correct (r) nucleotides, $f_{ins} = v(w)/v(r)$, can be deduced by measuring the kinetics of insertion of incorrect and correct nucleotides separately and then comparing the k_{cat}/K_m ratio for each (Fersht, 1985). When misinsertion kinetics are measured using the gel assay, the misinsertion efficiency is given by the relative insertion efficiencies of wrong and right in terms of $(I_i/I_{i-1})_{max}$,

$$f_{ins} = \frac{(k_{cat}/K_m)_w}{(k_{cat}/K_m)_r} = \frac{[(I_i/I_{i-1})_{max}/K_m]_w}{[(I_i/I_{i-1})_{max}/K_m]_r} \quad (6)$$

Note that it is not necessary to determine the enzyme dissociation rate, k_{off}, in order to calculate misinsertion efficiencies since the dissociation rate at site $i-1$ is the same regardless of the identity of the dNTP to be inserted at site i.

Misinsertion efficiencies next to different 5′-nearest neighbors for both KF$^-$ mutants were calculated by using eq 6 (Tables II and III). The greatest effect on f_{ins} of varying nearest neighbor primer termini in otherwise identical primer/template sequences was observed for the KF$^-$ (D424A) mutant polymerase, where AP insertion efficiencies were 0.31 ± 0.06, 0.23 ± 0.02, and 0.15 ± 0.03 next to G, C, and T, respectively. A change in the template base to the 5′-side of the target T site from A to G (Chart II) had little or no effect on the AP insertion efficiency next to primer T (Table II). Thus, as expected, the template cross-stacking perturbation appears to have less influence than the nearest neighbor interaction between the primer terminus and the incoming dNTP on the AP nucleotide insertion efficiency. There is a smaller difference in misinsertion efficiencies for KF$^-$ (D355A, E357A) on the different primer/templates. Misinsertion

[2] In cases where there is extension beyond the target site, the intensities of the bands extended to the target length and to greater than the target length must be summed ($(I_i + I_{i+1} + I_{i+2} + \ldots)/I_{i-1}$) to determine the relative rates of insertion at site i to polymerase dissociation at site $i-1$.

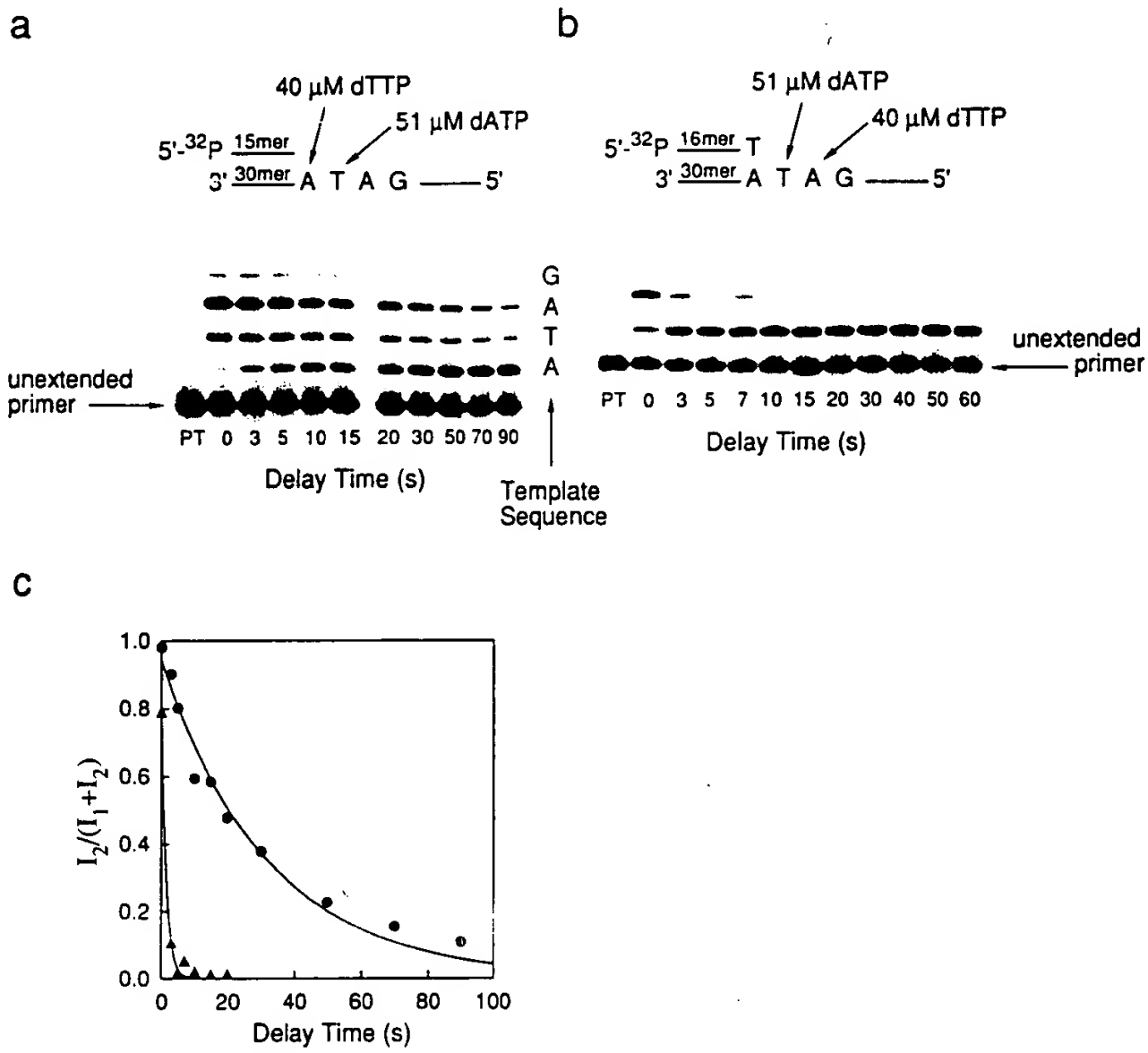


FIGURE 4: Measurements of k_{off} at two different sites on the same template, tA_1. (a) A gel showing measurement of KF⁻ (D424A) dissociation rate from up/tA_1 (15/30mer) after insertion of dTMP at 20 °C as described in Methods. Reaction mixtures contained 23 nM KF⁻ (D424A), 50 nM up/tA_1, 1 mg/mL trap DNA, 40 μM dTTP, and 51 μM dATP. (b) A gel showing dissociation of KF⁻ (D424A) following insertion of dAMP on pT/tA_1 (16/30mer). Reactions mixtures contained 23 nM KF⁻ (D424A), 50 nM pT/tA_1, 1 mg/mL trap DNA, 40 μM dTTP, and 51 μM dATP. (c) Plots of $I_2/(I_1 + I_2)$ versus time for the data in (a) and (b). For (a) the rate of dissociation following insertion of dTMP opposite A to form a 16mer is calculated where I_1 represents the intensity of the band due to the 16mer product and I_2 represents the sum of the band intensities for the 17mer and 18mer products. For (b) the dissociation rate following insertion of dAMP opposite T to form a 17mer is calculated where I_1 represents the intensity of the band due to the 17mer product and I_2 represents the intensity of the band due to the 18mer product. The plots decay as first-order exponentials from which k_{off} values of 0.031 s^{-1} for the 16/30mer (●) and >0.6 s^{-1} for the 17/30mer (▲) were calculated.

efficiencies next to nearest neighbor G, C, and T are 0.20 ± 0.02, 0.18, and 0.16, respectively.

Direct Competition Assay. We verified that the misinsertion efficiency deduced by measuring insertion kinetics for dAPMP and dAMP separately (eq 6) is equal to that obtained when dAPTP and dATP compete directly for insertion opposite T. Oligonucleotide products containing either 3′-terminus A or AP were clearly resolved by PAGE (Figure 5). The misinsertion frequency of 0.23 ± 0.01 found for direct competition was comparable to the misinsertion efficiency of 0.20 ± 0.02 determined by measuring kinetics separately.

Increase in Primer Utilization and Nucleotide Stimulation of Polymerase Dissociation. An unexpected observation was made concerning the extension kinetics of KF⁻ in the presence of dAPTP. When the gel assay was carried out in the absence of a trap under steady-state conditions using a large excess of DNA over polymerase and where each polymerase molecule must extend more than one primer/template, the rate of primer utilization increased as the concentration of dAPTP increased even though the concentration of the running-start base remained constant. This effect was especially pronounced for reactions with primers terminating in T (up/tA_1 and up/tA_2) and occurred to a smaller extent on the primer terminating in G (up/tC), but it was not observed on the primer terminating in C (up/tG). For example, during a 2-min reaction using up/tA_1 with 0.073 μM dAPTP about 5% of the primers were

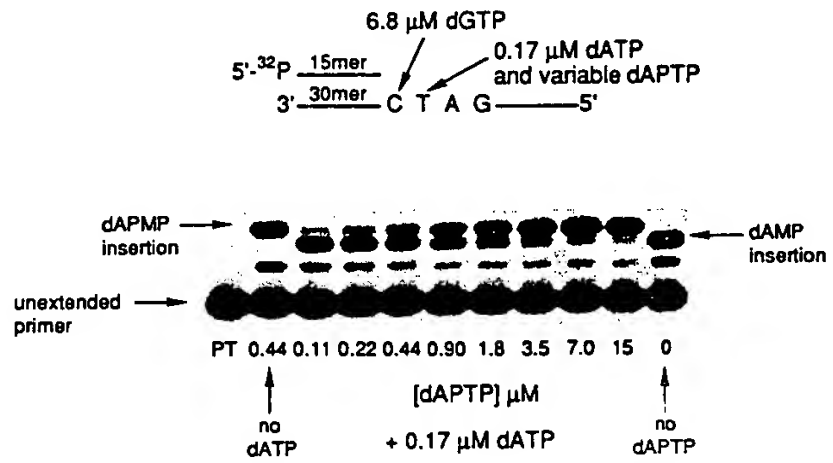


FIGURE 5: A gel showing the results of direct competition for insertion of dAMP and dAPMP into primer/template up/tC. Reaction mixtures contained a constant concentration (0.17 μM) of dATP and increasing concentrations of dAPTP as described in Methods. 5′-32P-Labeled 17mer products containing either a 3′-terminal A or AP were resolved by PAGE.

extended, but with 19 μM dAPTP about 30% of the primers were extended. The percentage of extended primers at intermediate concentrations increased as the concentration of dAPTP increased. These results imply that increasing the concentration of dAPTP, which is the second nucleotide to be inserted, increases the steady-state rate of primer utilization. It is not clear how the dAPTP is stimulating the rate of primer utilization by KF⁻ in an apparently sequence-dependent

manner. A similar observation was made for polymerase reactions with other mutants of Klenow fragment (Polesky *et al.*, 1992) and was attributed to dNTP-assisted polymerase–DNA dissociation.

Although product dissociation is not the rate-determining step in the extension of a single primer from site i – 1 to site *i*, it may be the rate-determining step in the overall steady-state cycle where a polymerase binds a primer/template, extends the primer, dissociates from that primer/template either before or after extension to site *i*, and then binds a *new* primer/template to continue the cycle. Dissociation rates for KF⁻ (D355A, E357A) bound to up/tA$_1$ were measured in the presence and absence of a ribonucleotide (GTP), which was not incorporated, to determine if high nucleotide concentrations could stimulate dissociation. Dissociation rates increased as the concentration of GTP increased. Values of k_{off} were 0.023, 0.039, and 0.065 s^{-1} for reactions which contained 0, 60, and 250 μM GTP, respectively. An increase in nucleoside triphosphate concentrations can apparently increase polymerase–DNA dissociation rates for these mutant polymerases; however, these nucleotide concentrations, which result in a 3-fold increase in the dissociation rate, are relatively high in comparison with the nucleotide concentrations used in kinetic assays.

Dissociation rates at site *i* – 1 and site *i* were also measured. During kinetics reactions, as the concentration of the target nucleotide is increased, the fraction of polymerase that dissociates at the site before the target site (site *i* – 1) decreases. If dissociation after insertion of the running-start nucleotide is significantly slower than dissociation after insertion of the target nucleotide, and if dissociation is the rate limiting step for steady-state primer utilization, as shown previously for insertion of correct dNTPs (Kuchta *et al.*, 1987), then the observed recycling time for polymerase will decrease as the fraction of polymerases dissociating after insertion of the target nucleotide increases. Thus, the rate of primer utilization will increase as the concentration of target nucleotide increases. Dissociation was measured on template tA$_1$ after insertion of dTMP and after insertion of dAMP or dAPMP (see above and Figure 4) and on tC after insertion of dGMP and after insertion of dAMP or dAPMP (data not shown). In both cases, dissociation was slower following insertion of the running-start base than following insertion of dAMP or dAPMP, which was too fast to measure. We estimate that dissociation rates following insertion of dAPMP on primer/templates pG/tC and pT/tA$_1$ by KF⁻ (D424A) must be greater than or equal to 0.7 and 0.5 s^{-1} for pT/tA$_1$ and pG/tC, respectively (assuming that reactions have progressed through 4 half-lives by the first time points at 4 and 5 s, respectively).

Both dNTP-dependent polymerase–DNA dissociation rates and different dissociation rates at different primer/template sites may be contributing to the observed increase in primer utilization. We have not observed a nucleotide stimulation effect with other polymerases, such as Sequenase (T7 DNA polymerase), avian myeloblastosis reverse transcriptase, or a 3′→5′ exonuclease-deficient mutant of T4 DNA polymerase, in similar assays (L. B. Bloom and M. F. Goodman, unpublished results).

DISCUSSION

Local DNA Sequence Effects on Insertion Kinetics. A main goal of this paper was to examine specific sequence effects on DNA polymerase reactions. Changing the identity of the base pair immediately 5′ to the nucleotide to be inserted had a significant effect on both insertion kinetics and misinsertion efficiencies. Because misinsertion efficiencies for AP are much larger than for natural nucleotides, it is a more sensitive probe for measuring differences in insertion due to differences in local DNA sequences. For KF⁻(D424A), we observed that misinsertion of dAPMP opposite T decreased to 0.31 ± 0.06, 0.23 ± 0.02, and 0.15 ± 0.03 for primers terminating with G, C, and T, respectively (Table II). Changing a downstream template base from A to G had a negligible effect on dAPMP insertion next to primer T. For the double mutant KF⁻(D355A, E357A), differences in misinsertion efficiencies with different nearest neighbors showed a smaller variation, although the trend was similar (Table III). Misinsertion next to G (0.20 ± 0.02) is greater than misinsertion next to T (0.16). Misinsertion efficiencies for dAPMP by both KF⁻ mutants ranged between 0.12 and 0.31, which is higher than the average value of 0.15 reported for T4 DNA polymerase at other sites (Bessman *et al.*, 1974; Clayton *et al.*, 1979; Pless & Bessman, 1983). These differences in relative rates of insertion given by the ratios of k_{cat}/K_m for a given dNMP next to different base-stacking partners seem to arise predominantly from differences in the k_{cat} values, which vary by a factor of 7–8. Values for K_m differ by a factor of 2–3 for a given dNMP on different primer/templates.

Base-stacking interactions between the incoming nucleotide and the 3′-primer terminus are likely to be contributing to the observed changes in insertion kinetics. Melting temperatures (T_m's) have been calculated for DNA doublets which contain either A·T or AP·T base pairs (Petruska & Goodman, 1985). The relative values of T_m reflect the relative stabilities of the doublets, which are due largely to base-stacking interactions between the two base pairs. Calculated T_m's are 69.7, 46.4, 39.8, and 28.5 °C when the 5′-base-stacking partners for AP are G, C, A, and T, respectively, and 86.4, 54.7, 54.5, and 36.7 °C when the 5′-base-stacking partners for A are G, C, A, and T, respectively (Petruska & Goodman, 1985), indicating that relative base-stacking energies vary with the 5′-base-stacking partner in the order G > C > A > T. Quantum-yield measurements made on synthetic primer/templates with identical sequences to the products of insertion by KF⁻ are correlated with the calculated T_m's. Quantum yields increased as the 5′-nearest neighbor base-stacking partner to AP was changed as follows: G < C < A ≈ T. An increase in quantum yield reflects a decrease in the interactions between AP and its neighboring bases on the primer/template which quench AP fluorescence. These experimental results support the calculated T_m's and relative stabilities of the doublets. The relative rates of both dAPMP and dAMP insertion in primer/templates with different 5′-nearest neighbor base-stacking partners correlate with the melting temperatures of the DNA doublets and the relative quantum yields. As seen in Tables II and III, the relative rates of insertion of both dAPMP and dAMP vary with the 5′-nearest neighbor in the order G > C > T.

Local DNA sequence has a large effect on KF⁻–DNA dissociation rates as well as insertion kinetics. Sequence-dependent differences in dissociation rates have also been observed for KF⁺ (Kuchta *et al.*, 1987) and HIV-1 reverse transcriptase (Yu & Goodman, 1992). About a 7-fold difference in k_{off} was seen for KF⁻ (D424A) for a primer terminating in G (0.26 ± 0.02 s^{-1}) compared with a primer terminating in T (0.037 ± 0.006 s^{-1}), and about a 20-fold difference was seen for KF⁻ (D355A, E357A) for primers terminating in G (0.24 ± 0.02 s^{-1}) and T (0.011 s^{-1}). Changing the base in the single-stranded region of the primer/template immediately following the target site from A to T increases

k_{off} by about a factor of 2 for KF$^-$ (D424A). KF has been shown to bind over 5–8 base pairs of the duplex region of the primer/template (Allen et al., 1989; Guest *et al.*, 1991a) in addition to bases in the single-stranded region of the template, so it was initially surprising that a change in 1 base pair could result in a 7–20-fold change in dissociation rate. Changes in some bases in the DNA strand are likely to have a more significant effect than changes at other sites. Changes in the region close to the site where the new nucleotide is to be inserted and where binding interactions with the polymerase are likely to be stronger would be expected to have a greater effect on the dissociation rate than changes in other more distant sites. When these differences in dissociation rates are viewed in terms of free energy differences for the two reactions, the energy differences are relatively small, only about 1 kcal/mol.

Interpretation of Kinetic Constants Derived from the Gel Assay. When the gel assay is done in the presence of a DNA trap, reactions are the result of a single encounter of a DNA polymerase with a primer/template. The kinetic constants correspond to rate-limiting steps in extension of primer/templates from site $i-1$ to site i and not product release. It is possible that the rate-limiting steps differ for different nucleotides and at different primer/template sites so that k_{cat} and K_m measured for one nucleotide or at one site may represent different mechanistic steps than k_{cat} and K_m measured for another nucleotide or at a different site. Misinsertion frequencies, however, are robust. The ratio of k_{cat}/K_m [or $(I_i/I_{i-1})_{max}/K_m$] for insertion of an incorrect and a correct nucleotide at the same site is a measure of the relative efficiencies of insertion (Goodman *et al.*, 1993).

Misinsertion efficiencies can be measured by direct competition of an incorrect nucleotide in the presence of a correct nucleotide, or they can be calculated, as in eq 6, from the values of k_{cat}/K_m [or $(I_i/I_{i-1})_{max}/K_m$] measured for each nucleotide separately (Boosalis *et al.*, 1987; Fersht, 1985). Misinsertion efficiencies for natural nucleotides are often hard to measure in direct competition experiments because they tend to be on the order of 10^{-3}–10^{-5}. Often concentrations of dNTPs are required which inhibit polymerase since the correct nucleotide must be maintained at a concentration in excess of the DNA concentration and a large pool bias of the incorrect nucleotide is required for the incorrect nucleotide to compete effectively for insertion. We have demonstrated that misinsertion efficiencies deduced by measuring kinetics for insertion of dAMP and dAPMP separately using the gel assay agree within experimental error with misinsertion efficiencies measured in direct competition. The misinsertion efficiency on primer/template up/tC was 0.23 ± 0.01 for dAPMP measured in direct competition with dAMP and 0.20 ± 0.02 from separate kinetics measurements (Table III).

2-Aminopurine as a Fluorescent Probe. Two methods for measuring polymerase kinetics were used in this study, fluorescence spectroscopy combined with rapid-mixing stopped-flow techniques and a gel fidelity assay. We found that rates and kinetic constants determined using pre-steady-state kinetic measurements were similar to those obtained from steady-state measurements.

AP has a much higher fluorescence intensity when present as a nucleotide than when present in DNA. We have shown that a decrease in AP fluorescence can be used to follow the kinetics of nucleotide insertion by a DNA polymerase, while a corresponding increase in fluorescence can be used to follow exonucleolytic removal of AP from a 3′-primer terminus (L. B. Bloom, M. R. Otto, Goodman, and Beechem, unpublished results). Although not extensively investigated in this study, the fluorescence of AP is also sensitive to interactions with KF$^-$. There is a small increase in fluorescence when double-stranded DNA containing AP at the 3′-primer terminus is bound to KF$^-$ and a corresponding increase in the rotational correlation time for AP in the complex.

Fluorescence spectroscopy is a powerful tool for studying DNA polymerase catalyzed reactions because dynamic information can be obtained about these reactions in real time. In combination with a fluorescent nucleotide analog such as 2-aminopurine, fluorescence spectroscopy can be used to follow polymerase-catalyzed reactions on a millisecond time scale, corresponding to single-turnover events for nucleotide insertion and removal. Another advantage of using AP as a fluorescent probe is that it does not appear to perturb DNA structure when paired opposite T (Nordlund *et al.*, 1989; Sowers *et al.*, 1986), unlike nucleotide analogs containing bulky fluorescent moieties. 2-Aminopurine's steric properties would be expected to be similar to normal nucleotides in the polymerase active site. Because the fluorescence properties of AP are sensitive to its environment and because it forms a reasonably good base pair with T, it may be a useful probe for studying dynamic interactions within DNA (Guest *et al.*, 1991b; Nordlund *et al.*, 1989) and dynamic interactions between other DNA binding proteins and AP-containing DNA.

Comparison of Fluorescence Measurements and the Gel Assay. A direct comparison of dAPMP incorporation kinetics was made using the fluorescence assay on both steady-state and pre-steady-state time scales and using the gel assay (Tables I–III). Results from all three assays were similar for the primer/templates used to measure the kinetics of insertion of dAPMP. Values for k_{cat}/K_m which correspond to the efficiency of insertion of 2-aminopurine ranged from 0.46 to 2.0 μM^{-1} s^{-1} (Table I) for the fluorescence assay under steady-state conditions and from 0.52 to 2.3 μM^{-1} s^{-1} for the gel assay (Table III). Values for k_{cat}/K_m measured under pre-steady-state conditions on primer/templates pC/tG and pT/tA$_1$ were 0.60 and 0.27 μM^{-1} s^{-1}, respectively, for KF$^-$ (D355A, E357A). Values of k_{cat} ranged from 0.41 to 2.0 s^{-1} for the fluorescence assay under steady-state conditions (Table I) and from 0.50 to 4.1 s^{-1} for the gel assay (Table III).

For the fluorescence assay, reactions were performed by preincubating the primer/template and enzyme and then initiating the reaction by the addition of dAPTP. If the first turnover of substrates to products occurred at a faster rate than subsequent turnovers, biphasic kinetics would have been observed in steady-state experiments where the ratio of enzyme to DNA was 1:14. Biphasic kinetics would indicate that a slower kinetic step occurred after incorporation of dAPMP that limited the steady-state rate of incorporation. We have measured the rate of insertion of dAPMP and the rate of dissociation following insertion of dAPMP. Insertion rates are relatively slow for dAPMP ($0.4 < k_{cat} < 5.0$ s^{-1} depending on sequence; Tables I–III), and dissociation rates are relatively high after insertion of dAPMP (estimated $k_{off} \geq 0.5$ s^{-1}), in comparison with rates of insertion and dissociation reported for the correct nucleotide (Kuchta *et al.*, 1987). The slow rate of insertion of dAPMP and the fast rate of dissociation are consistent with the apparent lack of biphasic kinetics and suggest that dissociation occurs at a rate similar to or faster than insertion, in contrast to the ≈800-fold greater rate of insertion than dissociation for the correct nucleotide (Kuchta *et al.*, 1987). If a small burst was obscured by noise in the data in experiments where a polymerase-to-DNA ratio of 1:14 was used, then we would have measured faster rates of insertion

in pre-steady-state experiments wh. enzyme-to-DNA ratios were higher (2:5 and 2:1). The similarity of the reaction rates measured in both pre-steady-state and steady-state experiments and the apparent lack of biphasic reaction kinetics suggest that both assays are measuring the same kinetic steps. Because the gel assay, done in the presence of a DNA trap, is the result of a single encounter of polymerase and DNA, all three assays are likely to be measuring steps in the reaction pathway that occur *prior* to release of product DNA (Goodman *et al.*, 1993). Mechanistic possibilities for this step include but are not limited to a conformational change in the enzyme, translocation along the DNA primer/template, and phosphodiester bond formation.

CONCLUSIONS

We have used two independent methods to measure insertion kinetics and fidelity for KF$^-$ as a function of the 5′-nearest neighbor base-stacking partner. Previous kinetic studies have not examined in a systematic manner the sequence dependence of DNA polymerase nucleotide insertion rates. While the effects of 5′-nearest neighbor base composition were examined in only a single sequence context, it is clear that these small changes cause significant variations in insertion kinetics, fidelities, and dissociation rates for KF$^-$. It remains to be seen whether these results hold within other sequence contexts and using other polymerases. The fidelity of insertion by different polymerases may be affected differently and to varying degrees by nearest neighbor base-stacking partners and different local sequence contexts. Incorporation efficiencies of dAPMP by bacteriophage T4 antimutator L141 DNA polymerase have been shown to vary greatly at different sites, while those for wild-type T4 DNA polymerase and KF$^+$ vary to a smaller extent at those same sites (Pless & Bessman, 1983). Polymerases with different biological roles and from different organisms may use different mechanisms and kinetic steps to achieve fidelity of nucleotide insertion (Capson *et al.*, 1992; Kati *et al.*, 1992; Kuchta *et al.*, 1987, 1988; Patel *et al.*, 1991; Wong *et al.*, 1991). Relatively small differences in base-stacking energies can lead to relatively large differences in k_{cat}/K_m at different sites since $\Delta\Delta G = -RT \ln[(k_{cat}/K_m)_{site\ A}/(k_{cat}/K_m)_{site\ B}]$. It will be important to determine how local DNA sequence generally affects insertion kinetics and fidelity for DNA polymerases from different sources.

A novel method for studying polymerase-catalyzed reactions in real time which is based on fluorescence changes for 2-aminopurine in different environments has been introduced. The validity of this spectroscopic methodology has been shown by the similarity of results obtained by using a kinetic gel assay. The real-time spectroscopic approach, however, can provide additional information concerning intermediate states not observable by examination of product formation alone. Although not presented in this study, dynamic interactions between a polymerase and a 2-aminopurine substrate (either nucleotide or DNA containing AP) in the active site can be examined during the reaction. In this paper, we have shown that fluorescence quench of 2-aminopurine on a millisecond time scale can be used to measure nucleotide insertion kinetics, and depolarization of 2-aminopurine can be detected on a nanosecond time scale, corresponding to rotational diffusion, in the active cleft of Klenow fragment. This methodology can be applied to studies of a wide range of enzymes, such as other DNA and RNA polymerases, exonucleases (the increase in AP fluorescence can be followed as AP is excised), and polymerase accessory proteins, to measure reaction kinetics on a millisecond scale and dynamic interactions between the enzymes and the AP substrates.

REFERENCES

Allen, D. J., Darke, P. L., & Benkovic, S. J. (1989) *Biochemistry 28*, 4601–4607.

Bessman, M. J., Muzyczka, N., Goodman, M. F., & Schnaar, R. L. (1974) *J. Mol. Biol. 88*, 409–421.

Boosalis, M. S., Petruska, J., & Goodman, M. F. (1987) *J. Biol. Chem. 262*, 14689–14696.

Capson, T. L., Peliska, J. A., Kaboord, B. F., Frey, M. W., Lively, C., Dahlberg, M., & Benkovic, S. J. (1992) *Biochemistry 31*, 10984–10994.

Clayton, L. K., Goodman, M. F., Branscomb, E. W., & Galas, D. J. (1979) *J. Biol. Chem. 254*, 1902–1912.

Derbyshire, V., Freemont, P. S., Sanderson, M. R., Beese, L., Friedman, J. M., Joyce, C. M., & Steitz, T. A. (1988) *Science 240*, 199–201.

Derbyshire, V., Grindley, N. D. F., & Joyce, C. M. (1991) *EMBO J. 10*, 17–24.

Echols, H., & Goodman, M. F. (1991) *Annu. Rev. Biochem. 60*, 477–511.

Eritja, R. E., Kaplan, B. E., Mhaskar, D., Sowers, L. C., Petruska, J., & Goodman, M. F. (1986) *Nucleic Acids Res. 14*, 5869–5884.

Fersht, A. R. (1985) in *Enzyme Structure and Mechanism*, W. H. Freeman and Co., New York.

Goodman, M. F., Creighton, S., Bloom, L. B., & Petruska, J. (1993) *Crit. Rev. Biochem. Mol. Biol. 28*, 83–126.

Guest, C. R., Hochstrasser, R. A., Dupuy, C. G., Allen, D. J., Benkovic, S. J., & Millar, D. P. (1991a) *Biochemistry 30*, 8759–8770.

Guest, C. R., Hochstrasser, R. A., Sowers, L. C., & Millar, D. P. (1991b) *Biochemistry 30*, 3271–3279.

Joyce, C. M. (1989) *J. Biol. Chem. 264*, 10858–10866.

Joyce, C. M., Chen Sun, X., & Grindley, N. G. (1992) *J. Biol. Chem. 267*, 24485–24500.

Joyce, C. M., & Grindley, N. D. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1830–1834.

Kati, W. M., Johnson, K. A., Jerva, L. F., & Anderson, K. S. (1992) *J. Biol. Chem. 267*, 25988–25997.

Kuchta, R. D., Benkovic, P., & Benkovic, S. J. (1988) *Biochemistry 27*, 6716–6725.

Kuchta, R. D., Mizrahi, V., Benkovic, P. A., Johnson, K. A., & Benkovic, S. J. (1987) *Biochemistry 26*, 8410–8417.

Mendelman, L. V., Boosalis, M. S., Petruska, J., & Goodman, M. F. (1989) *J. Biol. Chem. 264*, 14415–14423.

Nordlund, T. M., Andersson, S., Nilsson, L., Rigler, R., Gråslund, A., & McLaughlin, L. W. (1989) *Biochemistry 28*, 9095–9103.

Patel, S. S., Wong, I., & Johnson, K. A. (1991) *Biochemistry 30*, 511–525.

Petruska, J., & Goodman, M. F. (1985) *J. Biol. Chem. 260*, 7533–7539.

Pless, R. C., & Bessman, M. J. (1983) *Biochemistry 22*, 4905–4915.

Polesky, A. H., Dahlberg, M. E., Benkovic, S. J., Grindley, N. D. F., & Joyce, C. M. (1992) *J. Biol. Chem. 267*, 8417–8428.

Setlow, P., Brutlag, D., & Kornberg, A. (1972) *J. Biol. Chem. 247*, 224–231.

Sowers, L. C., Fazakerley, G. V., Eritja, R., Kaplan, B. E., & Goodman, M. F. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5434–5438.

Ward, D. C., Reich, E., & Stryer, L. (1969) *J. Biol. Chem. 244*, 1228–1237.

Wong, I., Patel, S. S., & Johnson, K. A. (1991) *Biochemistry 30*, 526–537.

Yu, H., & Goodman, M. F. (1992) *J. Biol. Chem. 267*, 10888–10896.

The Nucleotide Analog 2-Aminopurine as a Spectroscopic Probe of Nucleotide Incorporation by the Klenow Fragment of *Escherichia coli* Polymerase I and Bacteriophage T4 DNA Polymerase†

Michelle West Frey,‡ Lawrence C. Sowers,§ David P. Millar,|| and Stephen J. Benkovic*,‡

Department of Chemistry, 152 Davey Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802, City of Hope National Medical Center, 1500 East Duarte Road, Duarte, California 91010, and Department of Molecular Biology, Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, California 92037

Received February 16, 1995; Revised Manuscript Received May 5, 1995®

ABSTRACT: The fluorescent properties and their sensitivity to the surrounding environment of the nucleotide analog 2-aminopurine (2-AP) have been well documented. In this paper we describe the use of 2-AP as a direct spectroscopic probe of the mechanism of nucleotide incorporation by *Escherichia coli* Pol I Klenow fragment (KF) and bacteriophage T4 DNA polymerase. The nucleotidyl transfer reaction may be monitored in real time by following the fluorescence of 2-AP, allowing the detection of transient intermediates along the reaction pathway that are inaccessible through traditional radioactive assays. Previous studies with Klenow fragment [Kuchta, R. D., Mizrahi, V., Benkovic, P. A., Johnson, K. A., & Benkovic, S. J. (1987) *Biochemistry 26*, 8410−8417] have revealed the presence of a nonchemical step prior to chemistry and have identified this conformational change as the rate-limiting step of correct nucleotide incorporation. During correct incorporation, phosphodiester bond formation occurs at a rate greater than the conformational change and has not been measured. However, during misinsertion, the rate of the chemical step becomes partially rate limiting and it becomes possible to detect both steps. We have successfully decoupled the chemical and conformational change steps for nucleotide insertion by KF using the misincorporation reaction, and we present direct spectroscopic evidence for an activated KF′-DNA-dNTP species following the conformational change step which features hydrogen bonding between the incoming and template bases. In addition, we have utilized these same experiments to demonstrate the existence of a similar nonchemical step in the mechanism of dNTP incorporation by bacteriophage T4 DNA polymerase. This study provides the first direct evidence of a conformational change for T4 polymerase and emphasizes the importance of this step in a general polymerase kinetic sequence.

DNA polymerases are a family of enzymes responsible for the faithful duplication of DNA, *in vivo*. Though individual polymerases differ in their size, structure, requirement of accessory proteins, and role in DNA replication, their fundamental purpose is to catalyze the nucleotidyl transfer reaction—the addition of dNTPs onto the end of the growing DNA chain. Polymerase-mediated primer extension has been the focus of several studies in recent years aimed at gaining insight into the high fidelity achieved by this group of enzymes. These studies include extensive kinetic characterization of the polymerization mechanism for *Escherichia coli* Pol I and Klenow fragment (McClure & Jovin, 1975; Bambara et al., 1976; Bryant et al., 1983; Mizrahi et al., 1985, 1986a,b; Kuchta et al., 1987, 1988; Eger et al., 1991; Dahlberg & Benkovic, 1991; Eger & Benkovic, 1992) and bacteriophage T4 and T7 DNA polymerases (Patel et al., 1991; Wong et al., 1991; Donlin et al., 1991; Capson et al., 1992). Recent reports from our laboratory (Kuchta et al., 1988; Dahlberg & Benkovic, 1991; Eger & Benkovic, 1992) have established a minimal kinetic mechanism governing both correct and incorrect nucleotide incorporation for the Klenow fragment of *E. coli* Pol I. A significant feature of the mechanism is the proposed conformational changes prior to and following the chemical step of incorporation. Evidence for these conformational changes was obtained through a series of rapid quench experiments utilizing traditional radioactive assays. However, this type of assay is limited by the fact that it only measures product formation and cannot directly detect the presence of transient intermediates.

In this paper we describe a fluorescence-based continuous assay utilizing the nucleotide analog 2-aminopurine (2-AP).[1] The relatively high intrinsic fluorescence of 2-AP coupled with its extreme sensitivity to the surrounding environment (Ward et al., 1969; Guest et al., 1991; Bloom et al., 1993, 1994; Hochstrasser et al., 1994; Raney et al., 1994) provide a unique look into discrete steps along the reaction pathway. The data presented clearly demonstrate the ability to directly detect the previously documented conformational change in Klenow fragment, and they provide the first evidence for the existence of such a step in the bacteriophage T4 DNA polymerase mechanism.

† Supported by NIH Grants GM13306 (S.J.B.), GM44060 (D.P.M.), and GM50351 (L.C.S.).

* To whom correspondence should be addressed.
‡ The Pennsylvania State University.
§ City of Hope National Medical Center.
|| Scripps Research Institute.

® Abstract published in *Advance ACS Abstracts*, June 15, 1995.

[1] Abbreviations: Pol I, *E. coli* polymerase I; KF, Klenow fragment of pol I; KF exo⁻, (D335A, E357A) exonuclease deficient mutant of KF; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid, sodium salt; dNTP, deoxynucleoside 5′-triphosphate; T4 exo⁻, D219A exonuclease deficient mutant of bacteriophage T4 DNA polymerase.

0006-2960/95/0434-9185$09.00/0

EXPERIMENTAL PROCEDURES

Materials. Radioactive nucleotides [α-^{32}P]dATP, [γ-^{32}P]-ATP, and [α-^{32}P]TTP were purchased from New England Nuclear. T4 polynucleotide kinase was supplied by United States Biochemical (USB). Unlabeled, ultrapure nucleotides were obtained from Pharmacia. The (*Sp*)-dATPαS was synthesized by Dr. Jin Tann Chen. All other materials were of the highest purity commercially available.

Klenow Fragment. The Klenow Fragment (KF) exo$^-$ (D355A,E357A) was purified according to published procedures (Derbyshire et al., 1988) and then further purified by column chromatography using a BioRex 70 (Bio-Rad) anion-exchange resin which had been equilibrated with 10 mM PIPES, pH 7.0, and 1 mM DTT (PD buffer). The protein was eluted with a 1-L linear gradient 0−1 M NaCl in PD buffer. Fractions containing KF exo$^-$ (determined by A_{278}) were combined, dialyzed against 50 mM HEPES, pH 7.4, and 1 mM DTT, diluted 1:1 with 100% ACS grade glycerol, and stored at −20 °C. The stock concentration of the KF exo$^-$ was determined by $\epsilon_{278} = 6.32 \times 10^4$ M^{-1} cm^{-1} (Setlow et al., 1972) and active site titration (see below). Both values were in close agreement.

T4 D219A Exonuclease-Deficient Polymerase. The exonuclease-deficient mutant of T4 DNA polymerase (D219A) was purified as described previously (Frey et al., 1993) and stored at −70 °C. The concentration of the enzyme stock was determined by active site titration (Capson et al., 1992).

Oligonucleotides. The 13-mer and 20-mer oligonucleotides (see Figure 1) were supplied by Operon Technologies. An identical 20-mer except for the substitution of 2-aminopurine (2-AP) for thymine at position 7 was synthesized using the phosphoramidite method and will be described elsewhere. All oligonucleotide single strands (ss) were subjected to Hoefer gel purification as previously described (Capson et al., 1992) except that single strands were suspended in 25 mM Tris−acetate, pH 7.5, after desalting. DNA duplexes (13/20-mer and 13/20-AP) were purified on 3-mm non-denaturing gels (20% acrylamide/1X TBE/no urea), recovered as previously described (Capson et al., 1992), and quantitated as described below.

Enzyme Assays. Klenow Fragment assays were carried out in 50 mM Tris-HCl, pH 7.5. T4 polymerase assays were done in a buffer system consisting of 50 mM Tris−acetate, pH 7.5, 60 mM KOAc, and 10 mM 2-mercaptoethanol. All reactions were carried out at 20 °C. The rapid quench experiments were performed on the instrument described by Johnson (1986). Fluorescence assays were done using an Applied Photophysics stopped flow spectrometer with an excitation wavelength of 310 nm and a band-pass of 10 nm. Emission was monitored by using a 330-nm cutoff filter. All concentrations are initial concentrations unless otherwise noted.

Gel Electrophoresis. For radioactive assays, a 10-μL aliquot was removed at each time point and combined with 10 μL of gel load buffer (90% deionized formamide, 1× TBE, 0.25% bromophenol blue, and 0.25% xylene cyanol). Samples (7.5 μL) were then separated on denaturing gels (20% acrylamide/8 M urea). Products were visualized and quantitated using the Molecular Dynamics PhosphorImager and ImageQuant software version 3.3.

Incorporation of dTTP by KF exo$^-$. The 13/20-AP substrate (0.75 μM) was preincubated with an excess of KF exo$^-$ (3 μM) in one syringe of the rapid quench device, and then mixed with an equal volume of a solution containing $MgCl_2$ (20 mM) and correct nucleotide, dTTP (80 μM). The mixture was allowed to react for varying amounts of time (5−2000 ms) before being quenched with EDTA (0.5 M, pH 8.0). Products were separated and visualized as described above, and the quantities of each were calculated by determining the relative amounts of 13-mer and 14-mer, applying a 25% correction factor (vide infra), and multiplying by the DNA concentration (μM).

Stopped Flow Fluorescence Assay of Polymerization by KF exo$^-$. Excess KF exo$^-$ (6 μM) was incubated with 13/20-AP substrate (1.5 μM), and the reaction was initiated by mixing with an equal volume of a solution containing $MgCl_2$ (20 mM) and dTTP (80 μM). Quenching of 2-AP fluorescence was seen as an increase in signal voltage and was converted to concentration of DNA product by the factor illustrated in the caption of Figure 3. Multiple time courses were averaged (4−6 runs) to ensure proper signal to noise.

Rapid Quench Assay To Monitor the Incorporation of dTTP by T4 D219A Polymerase. The concentrations used for this assay were identical to that used in the KF exo$^-$ assay. Polymerization was initiated by mixing equal volumes of the polymerase/DNA solution (E·D) and Mg^{2+}·dTTP solution. The reaction was terminated at various times by the addition of EDTA (0.5 M, pH 8.0). Products were analyzed and corrected as described above.

Polymerization on 13/20-AP Substrate by T4 D219A Polymerase Followed by Fluorescence. Conditions for this experiment were the same as those used in the KF exo$^-$ fluorescence assay. Reactions were initiated in the stopped flow instrument by mixing equal volumes of the E·D and Mg^{2+}·dTTP solutions. The data from seven runs were averaged, and the fluorescence signal was converted to DNA concentration by the factor described in Figure 3.

Stopped Flow Fluorescence Assay of the Misincorporation of dATP opposite Template 2-AP by KF exo$^-$. Excess KF exo$^-$ (6 or 8 μM) was preequilibrated with the 13/20-AP substrate (1.5 or 2 μM) in one syringe of the stopped flow device and pushed against an equal volume of $MgCl_2$ (20 mM) and incorrect nucleotide dATP (80 μM) from a second syringe. The data shown are an average of at least four consecutive runs.

Radioactive Gel Assay of Misincorporation of dATP by KF exo$^-$. In a total reaction volume of 220 μL, KF exo$^-$ (3 μM) was incubated with 13/20-AP (0.75 μM) in assay buffer. The reaction was initiated by the addition of $MgCl_2$·dATP (final concentrations of 10 mM and 40 μM, respectively), and the time course was followed by removing 5-μL aliquots and quenching into 5 μL of EDTA (0.5 M, pH 8.0) at variable time intervals. Load buffer (10 μL) was added to each time point aliquot, and the products were separated and quantitated as described above.

Misincorporation of dATP opposite Template 2-AP by T4 D219A Followed by Fluorescence. An excess of T4 D219A polymerase (6 μM) was preincubated with 13/20-AP substrate (1.5 μM), and the reaction was initiated in the stopped flow system by mixing with an equal volume of a solution containing $Mg(OAc)_2$ (20 mM) and dATP or dATPαS (1 mM).

Radioactive Assay of T4 D219A Polymerase Misincorporation opposite Template 2-AP. 13/20-AP (0.75 μM) and T4 D219A polymerase (3 μM) were preequilibrated in a total

reaction volume of 50 μL. The reaction was initiated by the addition of $Mg(OAc)_2$-dATP or $Mg(OAc)_2$-dATPαS (final concentrations of 10 mM and 500 μM, respectively), and the progress was followed by quenching 5-μL aliquots into 5 μL of EDTA (0.5 M, pH 8.0) at various times. Load buffer (10 μL) was added to each time point aliquot, and the products analyzed as stated above.

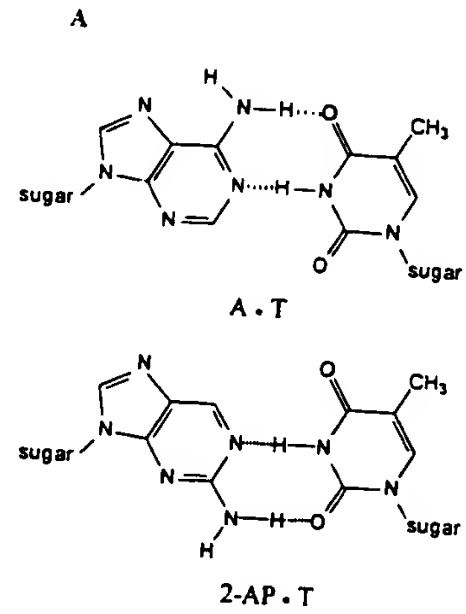
Determination of Active 13/20-AP Concentration. An accurate concentration of extendible ends was determined by the incorporation of [α-^{32}P]TMP under conditions of excess KF exo$^-$ (5 μM) over 13/20-AP substrate (varying concentrations), 10 μM dTTP (10 000 cpm/pmol), 5 mM $MgCl_2$, and 50 mM Tris-HCl, pH 7.5, in a total volume of 50 μL. Reactions were initiated by the addition of enzyme and were quenched after 10 or 60 s with either 0.1 M EDTA (20 μL) or phenol/chloroform (50 μL). Water (20 μL) was added to those reactions quenched by extraction. The amount of incorporated [α-^{32}P]TMP was determined by filter binding assay (Bryant et al., 1983), and the stock concentrations was calculated. To be sure that all 13/20-AP had been utilized, the extracted reaction samples (10 pmol) were 5′-^{32}P-end-labeled using standard protocols, separated on a 20% acrylamide denaturing gel, and visualized using the Molecular Dynamics PhosphorImager (data not shown). Through multiple trials it was found that approximately 25% of the 13-mer primer remained unextended up to 60 s.[2] Although the exact identity of this species is not known, it is not believed to interfere with the fluorescence experiments at hand because, upon conversion of the fluorescence data to DNA concentration, a full inventory of DNA product is obtained. However, the presence of this labeled, nonextendible 13-mer in the rapid quench assays interferes with the calculation of product formation (recall that the concentration of product is described by 14-mer cpm/(13-mer cpm + 14-mer cpm = rel%). Therefore, for the rapid quench assays, all data has been corrected for the unused 25% of 13-mer.

RESULTS

Kinetics of Incorporation of dTTP opposite Template 2-Aminopurine by KF exo$^-$. The DNA duplex structure remains relatively undisturbed when the nucleotide analog 2-aminopurine forms a Watson–Crick type base pair with thymine (Sowers et al., 1986). However, recent experiments by Bloom et al. (1993) suggested an alteration of the mechanism of incorporation by KF exo$^-$ when the 2-aminopurine 2′-deoxyribonucleoside 5′-triphosphate (2-APTP) was used (i.e., no burst of stoichiometric incorporation). Therefore, we wished to examine the kinetics of incorporation of dTTP opposite template 2-AP to determine what differences, if any, are manifest.

A 20-mer oligonucleotide with 2-AP at the 7th position was synthesized and annealed to a 13-mer primer, providing the primer/template (P/T) system shown in Figure 1. Initial rapid quench experiments to examine the pre-steady-state and steady-state rates of incorporation of a correct deoxynucleotide opposite template 2-AP by KF exo$^-$ revealed an initial burst phase followed by a second slower phase (M.

[2] In the misincorporation reactions followed by radioactivity, with time points > 10 min, the unused 13-mer primer described above was slowly utilized, perhaps as the polymerase slowly bypassed the misincorporated site (apurinic).



B

13/20-mer substrate:

```
5' - TCGCAGCCGTCCA
3' - AGCGTCGGCAGGTTCCCAAA
```

13/20-AP substrate:

```
5' - TCGCAGCCGTCCA
3' - AGCGTCGGCAGGTAPCCCAAA
```

FIGURE 1: (A) Watson–Crick type base pairing for the norm deoxyadenosine–thymidine and 2-aminopurine deoxynucleoside thymidine base pairs. (B) Duplex DNA sequences. The numbe indicate the lengths of the primer/template. The two substrat differ only in the substitution of the nucleotide analog 2-amin purine for the thymine at position 7 of the template strand (bold

W. Frey unpublished results) reminiscent of previous exper ments with the normal 13/20-mer substrate (Kuchta et al 1987; Dahlberg & Benkovic, 1991; Eger & Benkovic, 1992 To measure the burst rate of incorporation more accurately an experiment with excess enzyme over 13/20-AP substrat (single-turnover conditions) was done. A solution containing KF exo$^-$ (3 μM) and 13/20-AP (0.75 μM) in assay buffe was mixed with an equal volume of a solution containing $MgCl_2$ (20 mM) and dTTP (80 μM). The resulting time course is shown in Figure 2. The data were fit to a single exponential, and the burst rate constant was found to be on the order of 7.4 s^{-1} (Table 1).

Stopped Flow Fluorescence Assay of Incorporation of dTTP. The sensitivity of 2-AP to its environment, and its change in fluorescence upon going from single-stranded to duplex DNA have been well documented (Bloom et al., 1993, 1994; Ward et al., 1969; Guest et al., 1991; Hochstrasser et al., 1994; Raney et al., 1994). These properties, along with its similarity to the naturally occurring deoxyadenosine, foretell its use as a spectroscopic probe of polymerase action. The polymerization reaction was initiated in the stopped flow instrument by mixing equal quantities of an E·D solution (KF exo$^-$, 6 μM; 13/20-AP, 1.5 μM) and a Mg^{2+}-dTTP solution (20 mM $MgCl_2$; 80 μM dTTP). The reaction was monitored by the quenching of 2-AP fluorescence as the 13/20-AP was converted to 14/20-AP. The data shown in Figure 3 were converted to DNA concentration and fit to a single exponential, yielding a burst rate constant of 7.7 s^{-1} (Table 1). This value is in close agreement with that determined from the radioactive gel assay and indicates that the same step is being measured in both assays.

Table 1: Rate Constants Determined by Fluorescence or Radioactive Assays for Nucleotide Incorporation

enzyme	fluorescence assay (s^{-1}) 1st phase	fluorescence assay (s^{-1}) 2nd phase	radioactive assay (s^{-1})
KF exo$^-$			
correct incorporation		7.7 ± 0.1	7.4 ± 0.7
incorrect incorporation	0.17 ± 0.08	$0.0025 \pm 1.5 \times 10^{-4}$	$0.0025 \pm 2 \times 10^{-4}$
T4 exo$^-$			
correct incorporation		228 ± 5	220 ± 10
incorrect incorporation (dATP)		$0.02 \pm 6 \times 10^{-4}$	$0.021 \pm 5 \times 10^{-4}$
incorrect incorporation (dATPαS)	0.49 ± 0.02	$0.010 \pm 5.2 \times 10^{-4}$	$0.010 + 4 \times 10^{-4}$

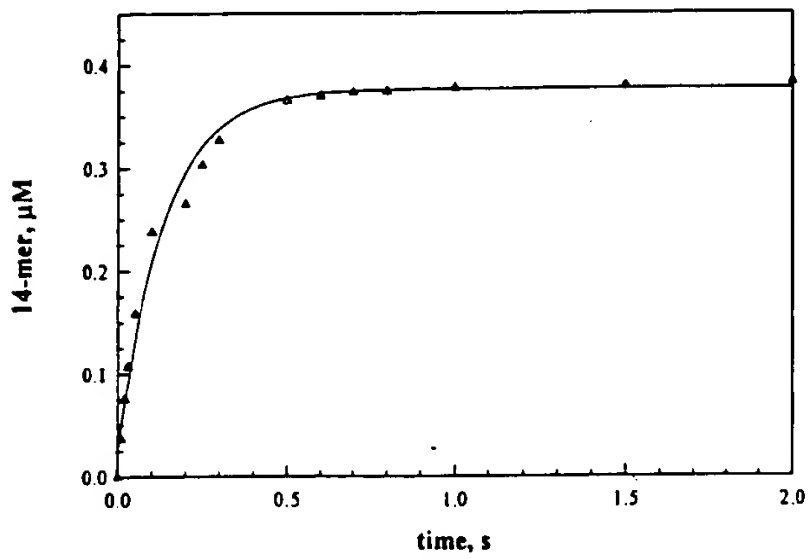


FIGURE 2: Incorporation of dTTP opposite template 2-AP as determined by rapid quench kinetics. Excess KF exo$^-$ (3 μM) was incubated with 13/20-AP (0.75 μM), and polymerization was initiated in the rapid quench instrument by mixing an aliquot of this solution with an equal volume of a solution containing $MgCl_2$ (20 mM) and dTTP (80 μM). The reaction was quenched at various times by the addition of 0.5 M EDTA, pH 8.0 (final concn = 0.35 M EDTA). Products were separated and analyzed as described in Experimental Procedures. Concentrations are initial concentrations unless otherwise noted. The data were fit to a single exponential with a rate constant of 7.4 s^{-1}.

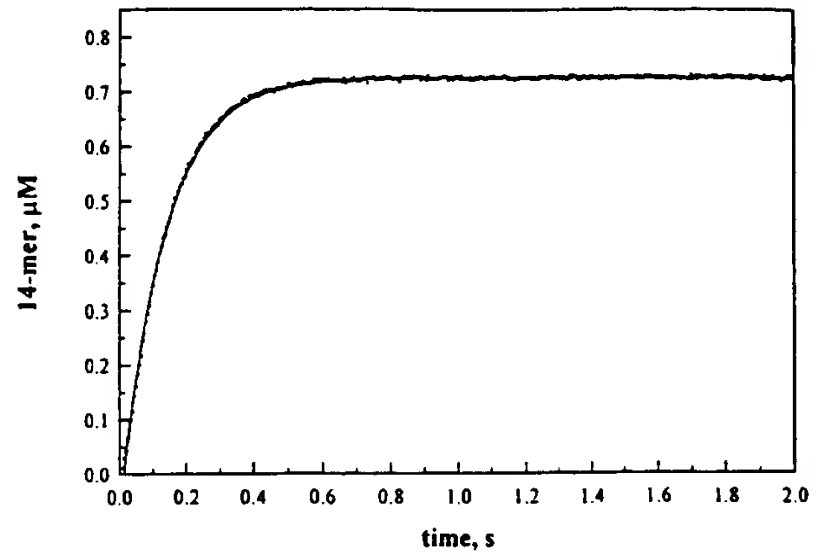


FIGURE 3: Stopped flow fluorescence assay of polymerization by KF exo$^-$. Excess KF exo$^-$ (6 μM) was preincubated with 1.5 μM 13/20-AP substrate (ED solution). Reactions were initiated by mixing an aliquot of the ED solution with an equal volume of a solution containing $MgCl_2$ (20 mM) and dTTP (80 μM). The data shown are an average of six consecutive runs and have been converted to DNA concentration using the factor DNA (μM) = $(\Phi + \Phi_0)/\Delta\Phi$XDNA (initial 13/20-mer DNA concentration), where Φ = fluorescence intensity at time t, Φ_0 is the initial fluorescence intensity, and $\Delta\Phi$ is the total change in fluoresence intensity. The data were fit to a single exponential, and the rate constant was determined to be 7.7 s^{-1} (Table 1). Unlike the radioactive assays, the DNA concentration is not corrected for the 25% unused material described in Experimental Procedures because it is not used by the polymerase and does not interfere with the calculation of product formation.

Rapid Quench Assay of dTTP Incorporation by T4 D219A pol. With the availability of the 13/20-AP substrate, another polymerase, T4 D219A pol (T4 exo$^-$), was examined in an attempt to further detail the kinetic mechanism of dNTP incorporation (see Discussion). Preliminary rapid quench experiments revealed biphasic kinetics for nucleotide incorporation opposite 2-AP by T4 exo$^-$ (M. W. Frey, unpublished results), similar to that reported previously (Frey et al., 1993). Conditions for the excess enzyme experiment were identical to those used in the KF exo$^-$ assay, and the time points were taken on the rapid quench instrument. The corrected data are shown in Figure 4. A computer fit to a single exponential provided a rate constant of 220 s^{-1}.

Fluorescence Assay of Polymerization by T4 D219A pol. Again, the conditions used for the KF exo$^-$ enzyme were applied to the T4 exo$^-$ and 13/20-AP experiment. T4 D219A (6 μM) was preincubated in the presence of 13/20-AP substrate (1.5 μM) in one syringe of the stopped flow instrument. The reaction was initiated by mixing an aliquot of the E·D solution with an equal volume of a Mg^{2+}·dTTP solution (20 and 80 μM, respectively) from a second syringe. The progress of the reaction was followed by the quenching of the 2-AP fluorescence, and the resulting change in Φ data converted to DNA concentration. The data and its fit to a single exponential are shown in Figure 5. The rate constant determined from the fluorescence assay, 228 s^{-1}, is in close agreement with that of the rapid quench gel electrophoresis assay.

Examination of the Misincorporation of dATP opposite Template 2-AP by Both Fluorescence and Radioactive Gel Assays. A tool for the further dissection of the polymerization mechanism is the misincorporation reaction. The misincorporation of dATP opposite template 2-AP in the 13/20-AP substrate by both polymerases (KF exo$^-$ and T4 D219A pol) was evaluated by both fluorescence and radioactive assays. The data from two representative time courses are shown in Figures 6 and 7. The raw fluorescence signal for the misincorporation of dATP opposite 2-AP by KF exo$^-$ is shown in the inset of Figure 6. An interesting feature of this time course is its apparent biphasic nature. The experiment, which is carried out under single-turnover conditions (excess enzyme over 13/20-AP substrate), is expected to fit a single exponential. However, the data are best described by a double exponential (inset, Figure 6). The identities of these two phases will be discussed below.

In contrast, the radioactive gel electrophoresis assay time course for the misincorporation of dATP opposite 2-AP by KF exo$^-$ (Figure 6, ●) is not biphasic, and it is best fit to a single exponential with a misincorporation rate constant of ca. 0.002 s^{-1} (Table 1). The second phase of the stopped flow fluorescence data which has been converted to DNA (μM) is overlayed with the gel assay data in Figure 6. The

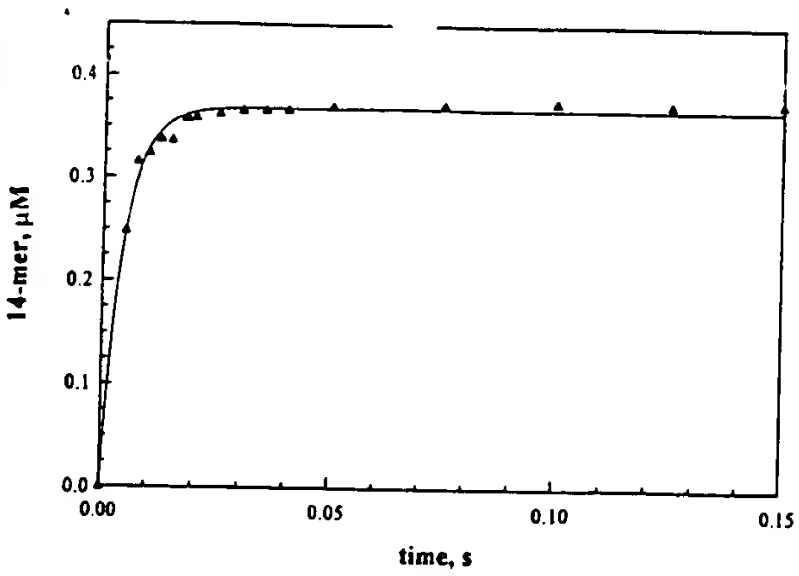


FIGURE 4: Rapid quench determination of the incorporation of dTTP by T4 D219A (T4 exo$^-$). Concentrations were identical with that used for the KF exo$^-$ (Figure 2). The reaction was initiated by mixing equal volumes of the two solutions and quenched at various times (5−150 ms) by the addition of 0.5 M EDTA, pH 8.0. (final concn = 0.35 M EDTA). Products were analyzed as described in Experimental Procedures. The resulting time course was fit to a single exponential, yielding a rate constant of 220 s^{-1} (Table 1).

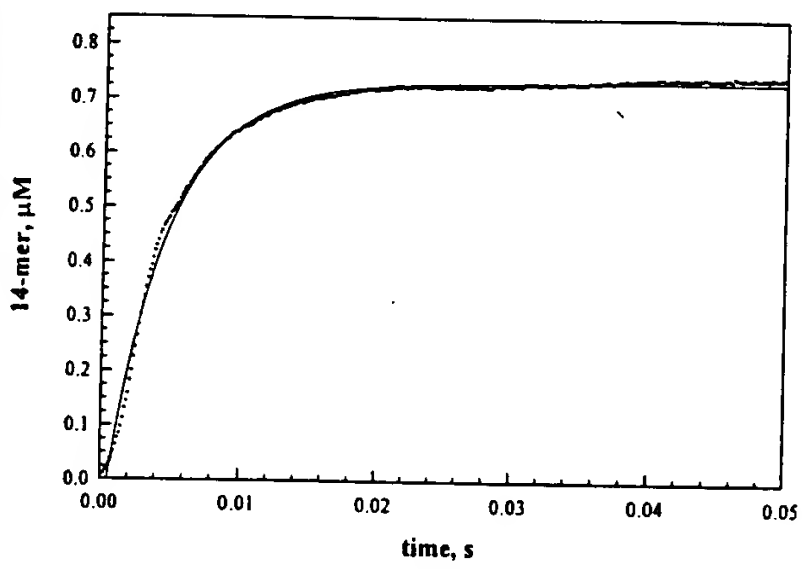


FIGURE 5: Fluorescence analysis of the polymerization by T4 D219A. Conditions were the same as those used for the KF exo$^-$ fluorescence assay. Reactions were initiated in the stopped flow by mixing equal volumes of an ED solution (6 μM T4D219A, 1.5 μM 13/20-AP) and a Mg^{2+}-dTTP solution (20 mM $Mg(OAc)_2$; 80 μM dTTP). The data shown are an average of seven runs, and fluorescence was converted to DNA (μM) by the factor illustrated in the caption of Figure 3. The data were fit to a single exponential, the rate constant of 228 s^{-1} is shown in Table 1.

data for both experiments are nearly identical, and the respective rate constants are equal within error (Table 1).

Similar experiments were carried out using the T4 exo$^-$ enzyme and the 13/20-AP DNA substrate. For those assays utilizing dATP, the reaction profile for both the fluorescence and radioactive assays (data not shown) is best described by a single exponential, unlike the results obtained for the KF exo$^-$ enzyme. However, when the α-phosphorothioate-substituted dATP is employed, the biphasic nature of the misincorporation reaction is again evident in the fluorescence-based assay (inset, Figure 7) but is absent from the radioactive assay (Figure 7, ●). The fluorescence data are best fit by a double exponential, and the exponential component describing the second phase overlays with the radiolabeled assay in Figure 7. The rate constants derived from the overlayed phases are nearly identical (0.01 s^{-1}) as shown in Table 1.

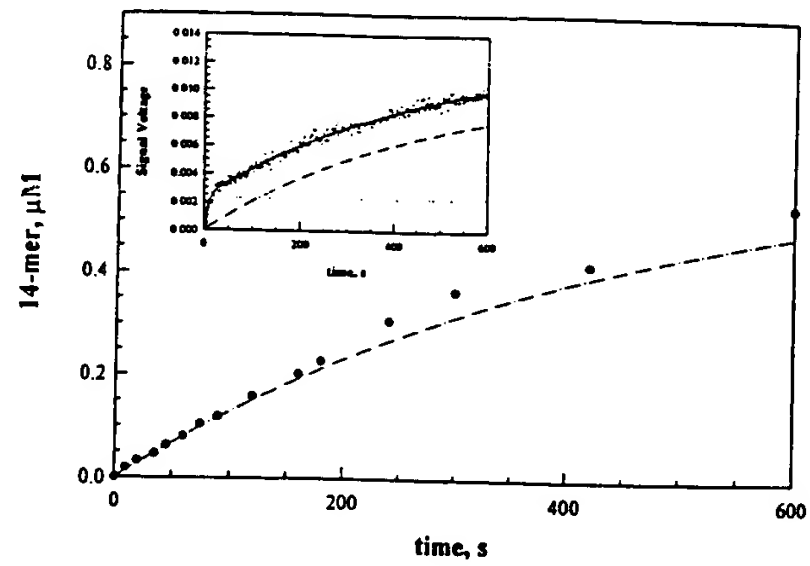


FIGURE 6: (Inset) Stopped-flow fluorescence assay of the misincorporation of dATP opposite template 2-AP by KF exo$^-$. Excess KF exo$^-$ (6 μM) was incubated with 13/20-AP substrate (1.5 μM) and pushed against an equal volume of Mg^{2+}-dATP solution (20 mM and 80 μM, respectively) in the stopped flow. The resulting time course was fit to the sum of two single exponentials. The identity of the initial burst is discussed in the text. The second phase of the fluorescence experiment was converted to DNA (μM) (•—•) and overlayed with the data from the radioactive gel assay data (●), indicating the second phase of the fluorescence assay time course is representative of product formation. Rate constants for both are on the order of 0.002 ± 0.0003 s^{-1} and are identical within error.

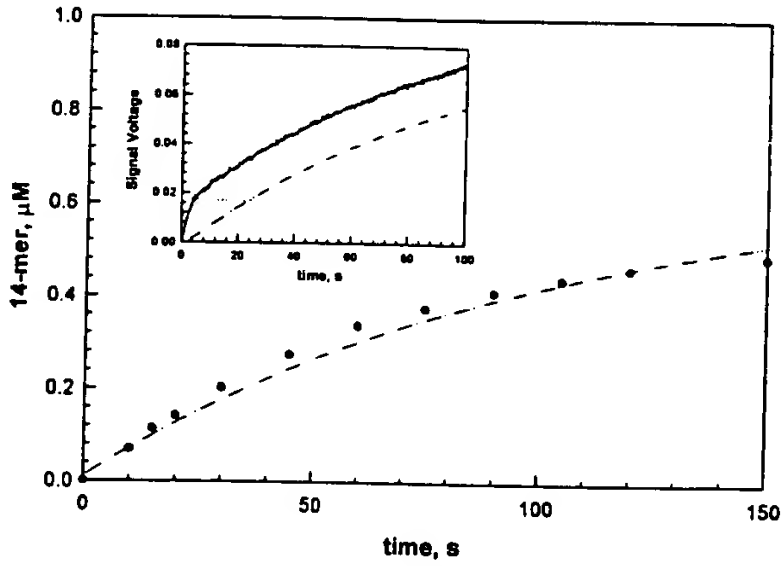
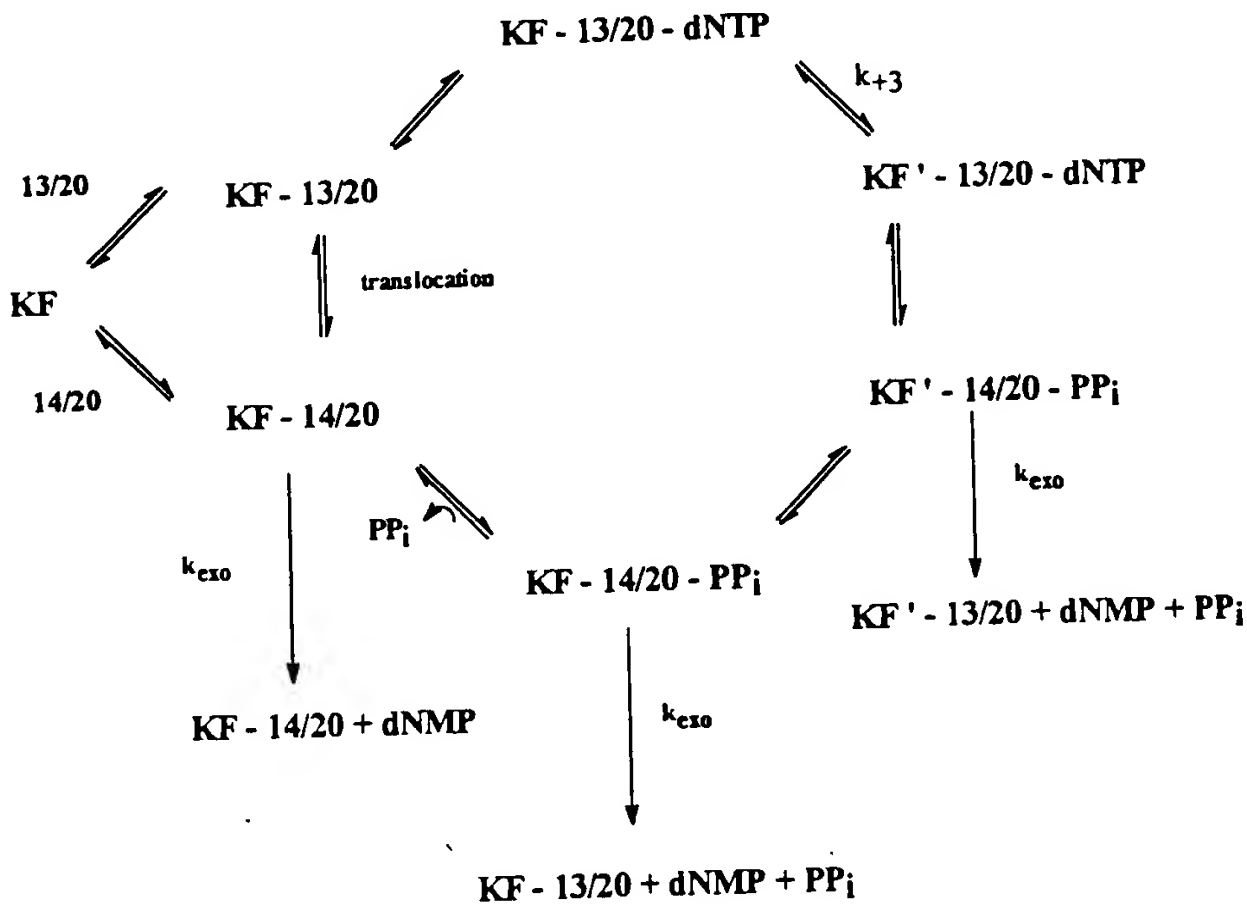


FIGURE 7: (Inset) Stopped flow fluorescence assay of the misinsertion of dATPαS opposite template 2-AP by T4 exo$^-$. Excess T4 exo$^-$ (6 μM) and 13/20-AP (1.5 μM) were preincubated in one syringe, and the reaction was initiated in the stopped flow by mixing with an equal volume of a Mg^{2+}-dATPαS solution from another syringe. The resulting biphasic time course was fit to the sum of two single-exponential functions. The second component of the fluorescence experiment was converted to DNA (μM) as described in the caption of Figure 3 (•—•) and is overlayed with the radioactive assay data (●). The rate constants derived for both are equal within error and are shown in Table 1.

DISCUSSION

The kinetic mechanism by which DNA polymerases carry out nucleotide incorporation has been the subject of intense study for the past decade [reviewed in Johnson (1993), Echols & Goodman (1991), Wang (1991), Young et al. (1992), Kornberg and Baker, (1992), McHenry (1991), Richardson et al. (1987), and Carroll and Benkovic (1990)]. Several researchers have attempted to elucidate the polymerization reaction pathway through the use of pre-steady-state and steady-state kinetic analyses which utilize radiolabeled substrates as a means of observing the progress of a given reaction. These experiments have provided great

Scheme 1



insight into the discrete steps along the polymerization pathway for several polymerases (Kuchta et al., 1988; Patel et al., 1991; Wong et al., 1991; Eger & Benkovic, 1992; Capson et al., 1992). However, there is a limitation to the radioactive gel assay, which only measures product formation and does not detect transient intermediates along the reaction pathway. In this paper we describe the use of the nucleotide analog 2-AP as a fluorescent probe of polymerase action. Previous reports have demonstrated that the kinetic mechanisms of nucleotide incorporation by KF exo$^-$ and T4 D219A polymerase (T4 exo$^-$) remain unchanged compared to the wild-type counterparts (Kuchta et al., 1988; Frey et al., 1993). The mutants were employed in the present study to avoid the complicating effects of exonuclease activity.

The DNA template strand was synthesized substituting 2-aminopurine 2′-deoxyribonucleoside at position 7, opposite the point of nucleotide insertion. The sensitivity of 2-AP fluorescence to its surrounding environment (Ward et al., 1969; Guest et al., 1991; Bloom et al., 1993, 1994; Hochstrasser et al., 1994; Raney et al., 1994) provides a spectroscopic handle with which to evaluate nucleotide incorporation. Since the fluorescence can be monitored continuously during the course of the reaction, the possibility exists of detecting intermediate species. Previous studies in our laboratory have been aimed at delineating the kinetic mechanism of nucleotide incorporation by the Klenow fragment of Pol I (Kuchta et al., 1987, 1988; Dahlberg & Benkovic, 1991; Eger & Benkovic, 1992) and bacteriophage T4 DNA polymerase (Capson et al., 1992; Frey et al., 1993). For the Klenow fragment, the minimal kinetic sequence is shown in Scheme 1. Of note is the rate-limiting step of processive synthesis, 50 s^{-1}, designated k_{+3}. This step has been assigned to a conformational change preceding chemistry for both correct (Kuchta et al., 1988; Dahlberg & Benkovic, 1991) and incorrect incorporation (Eger & Benkovic, 1992). Experiments presented here directly illustrate the presence of such a conformational change preceding the chemical step of phosphodiester bond formation, and permit comment on the nature of the KF′-13/20-AP-dNTP species.

Recent work by Bloom and co-workers described fluorescence-based pre-steady-state measurements of the insertion of 2-aminopurine 2′-deoxyribonucleoside 5′-triphosphate (dAPTP) into duplex DNA by Klenow fragment. However, no burst of nucleotide incorporation was observed in these experiments, making it difficult to interpret the origin of the fluorescence transients. In contrast, our preliminary experiments with 2-AP present in the template strand of the DNA substrate exhibited biphasic kinetics under conditions of limiting KF exo$^-$ (M. W. Frey, unpublished results). Studies were then undertaken to identify and measure the processes being observed by comparing the fluorescence and rapid quench kinetics assays.

The 13/20-AP substrate was used in the traditional radioactive rapid quench assay in order to compare the burst rate of nucleotide incorporation into the 13/20-AP substrate with that of the standard 13/20-mer (50 s^{-1}). Under conditions of excess KF exo$^-$ (to avoid the complicating slow step of DNA dissociation) the rate constant for dTTP (correct) incorporation was found to be ca. 7 s^{-1}, approximately 7-fold slower than that on the normal 13/20-mer. The stopped flow fluorescence assay, in which the quenching of the intrinsic fluorescence of the 2-AP moiety was monitored, yielded, within error, a similar rate constant. In view of previous evidence, it appears that the quenching of AP fluorescence is associated with a conformational change followed by a rapid phosphoryl bond formation.

Previous studies have shown that the 2-aminopurine is more soluble than the naturally occurring purines (Albert & Brown, 1954); 10 times more soluble than adenine, and 1000 times more soluble than guanine. The exclusion of water from the active site has been postulated (Petruska et al., 1988) to magnify the differences in free energy between correct and incorrect base pairs and to account for the high fidelity observed for most polymerases. The 7-fold decrease in the

observed rate here may reflect the exclusion of additional water molecules from the polymerase active site during incorporation opposite 2-AP. In turn, the conformational change may be responsible for the removal of water from the active site, and the presence of additional water slows the conformational change with respect to the 13/20-AP substrate.

To obtain further direct evidence for the proposed conformational change, a misincorporation reaction was employed. Experiments by Eger and Benkovic (1992) had previously demonstrated that the chemical step in the misincorporation process was slowed as compared to that in correct incorporation. For correct incorporation, chemistry occurs at a rate greater than the conformational change step, and the two processes are not distinguished. However, during misinsertion, where the rate of the chemical step is rate limiting, it may be possible to observe both chemistry and conformational change steps using both assay methods.

The misincorporation of dATP opposite template 2-AP was evaluated using both the fluorescence- and the radioactivity-based assays under conditions of excess enzyme. In the fluoresence experiment a double-exponential trace was acquired (Figure 6, inset). The radioactive gel electrophoresis assay, which measures the formation of 14(A)/20-AP mismatched products at all points along the reaction pathway (Figure 6, ●), is *not* biphasic, and it clearly matches the second component of the stopped flow fluorescence assay. Because no formation of a DNA product is associated with the initial phase found in the fluorescence assay, we believe this phase is representative of a change in the KF-13/20-AP fluorescence prior to product formation and after nucleotide binding. [Kinetic simulations were carried out using Scheme 1 and the rate constants reported previously (Eger & Benkovic, 1992) to demonstrate that the partial quenching of fluorescence associated with the initial phase was not due to simple collisional quenching by dATP of the ED species. This quenching would occur at the diffusion-limited rate of dATP binding, much faster than the observed rate of the initial fluorescence phase.] We therefore conclude that the partial quenching of 2-AP fluorescence observed in the misincorporation of dATP is indicative of an enzyme conformational change that aligns the incoming nucleotide and DNA substrate with partial hydrogen bonding in a position poised for phosphodiester bond formation. This configuration presumably facilitates interactions between the electronically excited AP chromophore and neighboring DNA bases, resulting in the observed fluorescence quenching. The rate constant for the initial phase (0.17 s^{-1}, Table 1) is somewhat slower than that measured for the conformational change during correct incorporation and may be explained by an overall slowing of the reaction as a result of the formation of an incorrectly base paired product.

Having demonstrated the ability to detect the previously documented conformational change of KF during nucleotide insertion, we attempted to identify a similar step for the bacteriophage T4 DNA polymerase. Previous attempts (Capson et al., 1992) to obtain support for a conformational change in the T4 DNA polymerase mechanism were unsuccessful. Experiments to examine the incorporation of a correct nucleotide into the 13/20-AP substrate were carried out with the T4 exo$^-$ as described for the KF exo$^-$. The fluorescence and radioactive gel assays yielded rate constants of 220 and 228 s^{-1}, respectively (Figures 4 and 5), a factor of 2 slower than that measured for the normal 13/20-mer. The misincorporation reaction was again utilized to decouple chemistry and a conformational change. The experiment was first attempted using dATP (data not shown). Unlike the results obtained with KF exo$^-$, the fluorescence assay was not biphasic and was best described by a single-exponential fit with a rate constant of 0.02 s^{-1}. The rate of misincorporated product formation is an order of magnitude faster than that measured for the KF exo$^-$, suggesting that the failure to detect a conformational change with T4 polymerase is due to the intrinsically higher rate of phosphodiester bond formation by this enzyme.

Table 2: Rates of Correct and Incorrect Nucleotide Incorporation by KF exo$^-$ and T4 Polymerase exo$^-$

enzyme	substrate	nucleotide	rate (s^{-1})	ratio
KF exo$^-$	13/20-mer (correct)	dATP	50[a]	
	9/20-mer (incorrect)	dATP	0.025[a]	2×10^3
	13/20-AP (correct)	dTTP	7.4	
	13/20-AP (incorrect)	dATP	0.0025	2.8×10^3
T4 exo$^-$	13/20-mer (correct)	dATP	400[b]	
	13/20-mer (correct)	dATPαS	200[b]	
	14/20-mer[c] (incorrect)	dATP	0.028	1.4×10^4
	13/20-AP (correct)	dTTP	228	
	13/20-AP (incorrect)	dATP	0.020	1.1×10^4
	13/20-AP (incorrect)	dATPαS	0.010	

[a] Data are from Eger and Benkovic (1992). [b] Data are from Capson et al. (1992). [c] The sequence of this 14/20-mer is

TGACGCACGTTGTC

ACTGCGTGCAACAGACTACG

(M. W. Frey, unpublished observations).

In an attempt to further slow the rate of chemistry, the phosphorothioate analog, dATPαS, was employed. The results shown in Figure 7 are very similar to those observed for the misincorporation by the KF exo$^-$. The fluorescence assay is again biphasic with the actual product formation (0.01 s^{-1}) being described by the second-exponential component, much like that seen for the KF exo$^-$, the 2-fold reduction in rate with dATPαS compared to dATP being the same as that previously determined for correct incorporation by T4 DNA polymerase (Capson et al., 1992). The initial phase of AP fluorescence quenching (0.49 s^{-1}, Table 1), is the first direct evidence for the existence of a conformational change in the kinetic mechanism of nucleotide incorporation by T4 DNA polymerase. Consequently, the kinetic step observed for the correct incorporation by T4 polymerase may also represent a conformational change prior to chemistry. A summary of the direct comparison of rate constants for correct and incorrect nucleotide incorporation on both substrates by both enzymes is presented (Table 2). The ratios for correct and incorrect insertion for both substrates are nearly identical, suggesting that no overall change in mechanism has occurred.

CONCLUSIONS

In this paper we have demonstrated the usefulness of the nucleotide analog 2-AP as a fluorescent probe of polymerase mechanism. A comparison of the ratios of correct vs. incorrect nucleotide incorporation on either 13/20-mer or 13/20-AP substrates for both polymerases indicates that no overall change in mechanism has occurred when 2-AP is substituted for thymine in the template. Rates measured with either the traditional radioactive assay or the fluorescence-

based assay are nearly identical. By monitoring the misinsertion reaction, we have detected the presence of an intermediate species that exists prior to phosphodiester bond formation, providing direct spectroscopic evidence for the proposed KF′-DNA-dNTP species. Furthermore, examination of T4 DNA polymerase by these methods has provided the first direct evidence of a previously undetected conformational change prior to chemistry and after nucleotide binding. The previous identification of a similar conformational change with the Klenow fragment (Eger & Benkovic, 1992; Dahlberg & Benkovic, 1991; Kuchta et al., 1988) as well as with T7 DNA polymerase (Wong et al., 1991) attests to the importance of this conformational change as a pivotal element in polymerase fidelity. This step, whether viewed as part of an induced fit mechanism for polymerase fidelity or as a necessary feature of nucleotide incorporation, acts as a discriminatory element in polymerase fidelity owing to its sensitivity to the nature of the nucleotide (complementary or noncomplementary to template strand). Differences in the rates of the conformational change step and in the chemical addition of nucleotides to the primer that have been observed for the KF exo$^-$ and T4 D219A polymerases may simply reflect the differing roles of these two polymerases in the cell, where the former serves as a repair enzyme and the latter as a replicative one.

REFERENCES

Alberts, A. & Brown, D. J. (1954) *J. Chem. Soc.*, 2060−2071.

Bambara, R. A., Uyemura, D., & Lehman, I. R. (1976) *J. Biol. Chem. 251*, 4090−4094.

Bloom, L. B., Otto, M. R., Beecham, J. M., & Goodman, M. F. (1993) *Biochemistry 32*, 11247−11258.

Bloom, L. B., Otto, M. R., Eritja, R., Reha-Krantz, L. J., Goodman, M. F., & Beechem, J. M. (1994) *Biochemistry 33*, 7576−7586.

Bryant, F. R., Johnson, K. A., & Benkovic, S. J. (1983) *Biochemistry 22*, 3537−3546.

Capson, T. L., Peliska, J. A., Kaboord, B. F., Frey, M. W., Lively, C., Dahlberg, M., & Benkovic, S. J. (1992) *Biochemistry 31*, 10984−10994.

Carroll, S. S., & Benkovic, S. J. (1990) *Chem. Rev. 90*, 1291−1307.

Dahlberg, M. E., & Benkovic, S. J. (1991) *Biochemistry 30*, 4835−4843.

Derbyshire, V., Freemont, P. S., Sanderson, M. R., Beese, L. S., Friedman, J. M., Steitz, T. A., & Joyce, C. M. (1988) *Science 240*, 199−201.

Donlin, M. J., Patel, S. S., & Johnson, K. A. (1991) *Biochemistry 30*, 538−546.

Echols, H., & Goodman, M. F. (1991) *Annu. Rev. Biochem. 60*, 477−511.

Eger, B. T., & Benkovic, S. J. (1992) *Biochemistry 31*, 9227−9236.

Eger, B. T., Kuchta, R. D., Carroll, S. S., Benkovic, P. A., Dahlberg, M. E., Joyce, C. M., & Benkovic, S. J. (1991) *Biochemistry 30*, 1441−1448.

Frey, M. W., Nossal, N. G., Capson, T. L., & Benkovic, S. J. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 2579−2583.

Guest, C. R., Hochstrasser, R. A., Sowers, L. C., & Millar, D. P. (1991) *Biochemistry 30*, 3271−3279.

Hochstrasser, R. A., Carver, T. E., Sowers, L. C., & Millar, D. P. (1994) *Biochemistry 33*, 11971−11979.

Johnson, K. A. (1986) *Methods Enzymol. 134*, 677−705.

Johnson, K. A. (1993) *Annu. Rev. Biochem. 62*, 685−713.

Kornberg, A., & Baker, T. A. (1992) *DNA Replication*, Freeman, New York.

Kuchta, R. D., Mizrahi, V., Benkovic, P. A., Johnson, K. A., & Benkovic, S. J. (1987) *Biochemistry 26*, 8410−8417.

Kuchta, R. D., Benkovic, P. A., & Benkovic, S. J. (1988) *Biochemistry 27*, 6716−6725.

McClure, W. R., & Jovin, T. M. (1976) *J. Biol. Chem. 250*, 4073−4080.

McHenry, C. S. (1991) *J. Biol. Chem. 266*, 19127−19130.

Mizrahi, V., Henrie, R. N., Marlier, J. F., Johnson, K. A., & Benkovic, S. J. (1985) *Biochemistry 24*, 4010−4018.

Mizrahi, V., Benkovic, P. A., & Benkovic, S. J. (1986a) *Proc. Natl. Acad. Sci. U.S.A. 83*, 231−235.

Mizrahi, V., Benkovic, P. A., & Benkovic, S. J. (1986b) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5769−5773.

Patel, S. S., Wong, I., & Johnson, K. A. (1991) *Biochemistry 30*, 511−525.

Petruska, J., Sowers, L. C., & Goodman, M. F. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1559−1562.

Raney, K. D., Sowers, L. C., Millar, D., & Benkovic, S. J. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 6644−6648.

Richardson, C. C., Beauchamp, B. B., Huber, H. E., Ikeda, R. A., Meyers, J. A. (1987) *DNA Replication and Recombination*, pp 151−171, Liss, New York.

Setlow, P., Brutlag, D., & Kornberg, A. (1972) *J. Biol. Chem. 247* 224−231.

Sowers, L. C., Fazackerly, G. V., Eritja, R., Kaplan, B. E., & Goodman, M. F. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5434−5438.

Wang, T. S.-F. (1991) *Annu. Rev. Biochem. 60*, 513−552.

Ward, D. C., Reich, E., & Stryer, L. (1969) *J. Biol. Chem. 244*, 1228−1237.

Wong, I., Patel, S. S., & Johnson, K. A. (1991) *Biochemistry 30*, 526−537.

Young, M. C., Reddy, M. K., & von Hippel, P. H. (1992) *Biochemistry 31*, 8675−8690.

BI9503530

ALEXANDROVA ET AL.

: and then with N-succinidyl
Twice modified dNTPs **Ia,b**
pounds were effective and
and δ, but were not
human DNA polymerase β

y-3',5'-diacetyluridine was
replacement of OH groups
n, the resulting nucleosides
reduction gave dNTPs **IIb**,
-6-aminohexanoate or FITS,

nates **IIa-e** demonstrated
s **IIIa-d** were substrates for
hosphonate, but they were
dUTPs, **IIIe-f**, the affinity
that for their counterparts
osition.
's **IVa-g** on their substrate
gation effectively catalyzed
e residue, whereas **IVg** was

ate groups in dNTP by
properties towards DNA
under investigation keep
xtend the primer.
ian Foundation of Basic
and Outstanding Schools,

Bioorg Chem. (Russian)

torova, L.S., Krayevsky,
n.

NUCLEOSIDES & NUCLEOTIDES, 18(4&5), 965–966 (1999)

REINVESTIGATION OF 4-THIOTHYMIDINE-5'-TRIPHOSPHATE SYNTHESIS

Bazin, H.[1]* and Sauvaigo, S.[2]
[1]CIS biointernational/ DIVT/ Research and New Technologies, BP 175, F-30203 Bagnols/Cèze Cedex, France. E-mail: hbazin@compuserve.com
[2]CEA Grenoble/DRFMC/SCIB/Lésions des Acides Nucléiques, 17 rue des Martyrs, F-38054 Grenoble Cedex 9, France.

ABSTRACT : The 4-Thiothymidine-5'-triphosphate **1** (S^4TTP) was known to be a substrate for polymerase, however a commercial sample of this compound failed to be incorporated into DNA. Mass spectrometry combined to alkaline phosphatase digestion and ^{31}P-NMR showed that this sample was in fact 5'-chloro-5'-deoxy-4-thiothymidine-3'-triphosphate **2**. The desired S^4TTP was synthesized by two alternate routes, was fully characterized and was shown to be incorporated in a DNA polymerase assay.

Sample analysis: The commercial dNTP (USB #77153, Amersham) gave the following MS-(ES^-), after exchange with TEA : $(M-H)^-$ 515.1 (100%) and 517.2 (40%). These data were inconsistent with a mass of 498 expected for S^4TTP (Calc. for free acid $C_{10}H_{17}N_2O_{13}P_3S$). The 18 units shift and the isotopic motive suggested the replacement of a hydroxyl by chlorine. The ^{31}P-NMR showed the presence of a triphosphate and the UV (λmax = 335nm in H_2O) was consistent[1] with S^4T structure. The dNTP (2.5 µl of the 10mM commercial solution) in 50µl buffer (0.1M Tris pH 9, 0.1M NaCl, 15 mM $MgCl_2$) was treated with 5µl Snake Venom Phosphodiesterase (Pharmacia, 5U/µl) no change was observed in HPLC[2] in contrast with dTTP which was hydrolysed to dTMP. The dNTP (10µl of 10mM) in 85µl buffer (as above) was treated with 5µl of alkaline phosphatase (125U), within 15mn HPLC[3] showed complete hydrolysis to a new product (Rt= 20mn)[3]. This product was distinct from authentic 4-thiothymidine[1] (Rt= 13.1 mn)[3] and the ES-MS data [$(M-H)^-$ 275.1 (50%), 277.2 (15%)] was consistent with chlorine substituted 4-thio-thymidine. The proton coupled ^{31}P-NMR of the dNTP sample displayed a quartet at -11.2

ppm (Pα) with J_{PH}= 8.2Hz and J_{PP}=19Hz. Upon selective decoupling of $H_{3'}$ (5.13ppm), the quartet collapsed to a doublet (J=19Hz) and upon $H_{5'}$, $_{5''}$ decoupling (4.1ppm), the signal remained unaffected. These data indicated a triphosphate chain at 3'position, correlated with a resistance to SVP and in favour of structure **2** for commercial dNTP.

4-STTP Synthesis: To understand the origin of compound **2**, and considering the reported S^4TTP synthesis[4], one striking feature was the use of pyridine/$(EtO)_3PO$ mixture for phosphorylation (presumably to prevent glycosyl cleavage). Since the S^4TTP used in ref.5 was prepared in $(EtO)_3PO$ alone, we tested the action of pyridine. S^4T [1] (50µmol) in pyridine (24µl)/$(EtO)_3PO$ (150µl) mixture (0°C) was reacted with $POCl_3$ (50µl) for 5h. The main product was eluted on DEAE-sepharose with TEAB 0.18M (Rt=19.5mn)[2], the ^1H coupled ^{31}P-NMR (doublet, J_{PH}=7.3Hz) and MS-(ES$^-$) [(M-H)$^-$ 355.1(100%), 357.1 (35%)] were compatible with 5'-chloro-5'-deoxy-S^4T-3'-phosphate structure. Similarly S^4T was phosphorylated in $(EtO)_3PO$ alone, the main product eluted on DEAE-sepharose with TEAB 0.17M (Rt=7.7mn)[2]. The ^{31}P-NMR and MS-(ES$^-$) [(M-H)$^-$ 337.1(100%)] were compatible with 4-thiothymidine-5'-monophosphate **3** structure. The S^4TMP **3** was converted to S^4TTP **1** by the carbonyl-diimidazole procedure[4], the product (8.7µmol) was eluted on DEAE-sepharose with TEAB 0.38M (Rt=16.3mn)[2], MS-(ES$^-$): (M-H)$^-$ 497.1. The S^4T was also phosphorylated by an alternate procedure[6] and yielded S^4TTP as above.

CONCLUSIONS: Pyridine induced the formation of 5'-chlorinated nucleotide in the $POCl_3$ phosphorylation, and reported S^4TTP data[4] were questionable. Authentic S^4TTP was synthesized and was incorporated in a polymerase assay (Pol I Klenow fragment).

REFERENCES

1. Kraszewski, A. ; Delort, A.M. and Teoule, R., *Tetrahedron Lett.*, **1986**, *27*, 861-864.
2. Lichrospher^R RP18E(5µ)125x4, Eluant A: 4% ACN in 5mM TBAP 50mM Na_2HPO_4 pH6; B : ACN. 10 to 30%B in 20min, then 30 to 50%B in 5mn. 1ml/mn.
3. Same column. Eluant A: TEAB 50mM pH7; B : ACN. 0 to 60%B in 30mn. 1ml/mn.
4. Scheit, K.H. and Faerber, P. *Nucleic Acids chemistry.* Townsend, L.B. and Tipson, R.S. (Ed), John Wiley and Sons, **1978** , Vol.2, 793-799.
5. Hofer, B and Köster H. *Nucleic Acids Res.*, **1981**, *9*, 753-767 and references herein.
6. Ruth, J.L and Cheng, Y-C. *Mol. Pharmacol.*, **1981**, *20*, 415-422.

AN *IN SITU* PIG LIV
TOOL FOR THE
PH

Carlo Ballatore,[1*] Ch

1. Welsh School of Pharm
3XF, UK. 2. Rega Institu

ABSTRACT: The pig liv
vitro the first step of the m
amino acyl phosphoramid
necessary but not *sufficie*
predictive tool for the like
Relationship establishment

Among the different ma
nucleosides, the phosphor
shown higher anti viral p
nucleosides[1]. These are ty
(So324)[2] **(FIG. 1)**. The r
monophosphate (NMP) di
- mediated first phosphor
The metabolic pathways t
possible enzymatic steps
hydrolysis of the carboxyl
N bond. The PLE assay h
involved in the generation
described as a key i
phosphoramidate is a goo

(i) Tables, should be numbered and should include a heading. When explanatory footnotes are included they should be typed at single-line spacing, and the table, together with its footnotes must not exceed the normal page length.
(j) Papers containing detailed photographs may be printed on smooth-coated paper. A request for printing on smooth-coated paper must be made at the time the manuscript is submitted, and a charge of $25 (£15) will be made.
(k) Footnotes, including abbreviations and changes of address of authors, should be indicated with superscript figures, and included in the references.

References

The citation of journals (abbreviated in the style of 'Chemical Abstracts'), books and multi-author books should conform with the following examples (without any further indenting).

1. Razzell,W.E. and Khorana,H.G. (1958) Biochim. Biophys. Acta 28, 562-566.
2. Davidson,J.N. (1969) The Biochemistry of the Nucleic Acids, 6th edn. pp. 177-178, Methuen, London.
3. Burdon,R.H. (1971) in Progress in Nucleic Acid Research and Molecular Biology, Davidson,I.N. and Cohn,W.E. Eds., Voll II, pp. 33-79, Academic Press, New York.

Single-line spacing should be used throughout, including between individual references.

Nomenclature

As far as possible, authors should follow the Recommendations of the IUPAC-IUB Commission on Biochemical Nomenclature, and particularly the abbreviations for nucleic acids, polynucleotides and their constituents 1971) J. Mol. Biol. 55, 299-305.

Reprints

A full list of pre-publication reprint prices can be found in the first January issue of the journal and a reprint order form will accompany our notification of acceptance of a manuscript. Thirty reprints are supplied free of charge.

Volume 9 Number 4 1981 Nucleic Acids Research

Enzymatic synthesis, ligation, and restriction of DNA containing deoxy-4-thiothymidine

Bernd Hofer and Hubert Köster

Institut für Organische Chemie und Biochemie der Universität Hamburg, Martin-Luther-King-Platz 6, D-2000 Hamburg 13, GFR

Received 5 January 1981

ABSTRACT

Phage fd RF I DNA[1] about 90% substituted by deoxy-4-thiothymidine (s^4T_d) in the codogenic strand was synthesized by the simultaneous actions of DNA polymerase I and DNA ligase. While the rate of DNA synthesis was considerably reduced, the yield was not affected in the presence of s^4T_dTP. The conversion of RF II to RF I DNA by DNA ligase was even improved. This effect seems to be related with an altered ratio of affinity of polymerase and ligase for the s^4T_d-containing substrate. The presence of the base analogue in the DNA was verified independently by chromatographic and spectroscopic methods. The modified genome could be cleaved by restriction endonucleases Hpa II $(C/CGG)_d$ and Taq I $(T/CGA)_d$. A number of the fragments produced showed altered mobilities under the conditions of polyacrylamide gel electrophoresis.

INTRODUCTION

Modified nucleic acids have proved to be useful tools for many purposes in molecular biology. For example, a large number of studies on nucleic acid-nucleic acid (e.g., 2-5) and nucleic acid-protein interactions (e.g., 6-15) have been published which made use of modified DNAs. Moreover site-directed mutagenesis can also be performed by the incorporation of nucleotide analogues into DNA (16) or RNA (17).

At present a limited number of modified nucleotides is available that can be enzymatically incorporated into nucleic acids of natural sequence. One of the positions in DNA not accessible to modification so far was the 4-keto group of thymidine (T_d), which is of particular interest as it is involved in Watson-Crick base pairing.

Enzymatic synthesis of alternating polymers containing deoxy-4-thiothymidine (s^4T_d) were reported several years ago

(18,19). But difficulties were met in attempts to incorporate the analogue into activated calf thymus DNA (18,20). Since such a DNA is no well-defined template and detailed data on the reaction were not available, we found it worthwhile to re-examine this subject.

We used E.coli DNA polymerase I, s^4T_dTP and a template of natural origin and known sequence (21), the circular single-stranded DNA of bacteriophage fd. In the presence of DNA ligase we were able to obtain the double-stranded closed circular RF I DNA containing the modified nucleotide in the minus-strand. The DNA or its degradation products were characterized by spectroscopic, electrophoretic and chromatographic methods.

MATERIALS AND METHODS

Enzymes. E.coli DNA polymerase I and T4 DNA ligase were isolated in this laboratory by H. Müller and R. Frank, respectively, according to standard methods (22,23). E.coli DNA ligase was prepared similar to published procedures (24,25). Specific activity of DNA polymerase I: 8160 units per mg. One unit catalyses the incorporation of 10 µmol of nucleotides into polyA_dT_d in 30 min at 37°C under assay conditions (22). One unit of DNA ligase converts 100 nmol (nucleotides) of poly($A_d \cdot T_d$) into exonuclease III-resistant covalently closed circles in 30 min at 30°C under assay conditions (23,26). Restriction endonuclease Taq I from Thermus aquaticus was a kind gift of Dr. H. Mayer, Stöckheim. Restriction endonuclease Hpa II from Haemophilus parainfluenzae was supplied by Miles Laboratories, Elkhart, Indiana, USA. Micrococcal nuclease from Staphylococcus aureus, spleen phosphodiesterase, DNase I from bovine pancreas (grade I), and snake venom phosphodiesterase were purchased from Boehringer, Mannheim, FRG.

Substrates. Viral single-stranded DNA from bacteriophage fd and fd-specific oligonucleotide primers were prepared as previously described (10). Nucleoside triphosphates were purchased from Boehringer, Mannheim, FRG. [^{3}H]C_dTP was supplied by New England Nuclear, Boston, Mass., USA. [α-^{32}P]A_dTP was from The Radiochemical Centre, Amersham, UK. s^4T_dTP was synthesized similar to the procedure published by Scheit (27). As judged from chromatographic analysis and UV absorption spectrum it was more than 95% pure.

DNA synthesis and ligation. These procedures were performed as described (15) with the following modifications: Either C_dTP ($1.6 \cdot 10^4$ cpm/nmol) or A_dTP ($9.2 \cdot 10^5$ cpm/nmol) were radiolabeled. 60 units of DNA polymerase I and 2.5 units (in the presence of T_dTP) or 0.9 units (in the presence of s^4T_dTP) of T4 or E.coli DNA ligase were used. The mixture was incubated at 20°C for 4 h (in the presence of T_dTP) or 16 h (in the presence of s^4T_dTP). After deproteinization preparative purification of RF I was performed by acid phenol extraction (28).

Chromatographic analysis of s^4T_dMP incorporation. DNA synthesis in the presence of [α-^{32}P]A_dTP was performed as described above. The ethanol precipitated pellet was washed with ethanol and dissolved in 5mM Tris base. The DNA was subsequently degraded by St. aureus nuclease and spleen phosphodiesterase as described by Josse et al. (29). The digests were chromatographed on silica gel thin layer (60 F254 from Merck, Darmstadt, FRG) using ethanol-1M ammoniumacetate, pH 7.5, 7:3 (v/v) as solvent. The products were detected by autoradiography on Kodirex X-ray films from Kodak.

UV-spectroscopic analysis of s^4T_dMP incorporation. Spectroscopic measurements were performed at room temperature in a Gilford 2400-S spectrophotometer. RF I DNA was dissolved in 100 mM Tris·HCl, pH 8.0, 10 mM $MgCl_2$, 0.2 mM EDTA. Degradation to mononucleotides was performed in this solution at room temperature by subsequent incubation with DNase I from bovine pancreas and snake venome phosphodiesterase (30).

Fragmentation of s^4T_d-substituted DNA. Hpa II: 2 µg of RF I DNA were incubated in 60 µl of 10 mM Tris·HCl, pH 7.5; 10 mM $MgCl_2$; 5 mM KCl; 1 mM DTE; 2.5% glycerol with 5.5 units of Hpa II at 37°C for 7 h. Taq I: 2 µg of RF I DNA were incubated as above in 10 mM Tris·HCl, pH 7.5; 10 mM $MgCl_2$; 5 mM KCl; 10 mM ME; 5% glycerol with 15 units of Taq I.

Gel electrophoresis. The conditions for agarose (buffer A) and polyacrylamide gel electrophoresis as well as visualization and photography of DNA bands habe been described earlier (10, 15). Buffer B for agarose gel electrophoresis contained 40 mM

Tris·HOAc; 5 mM NaOAc; 2 mM EDTA; pH 7.8 (31).

Velocity sedimentation. Velocity sedimentations were performed in alkaline 5-20% sucrose gradients. Runs were for 50 min at 50.000 rpm and 15°C in a Beckmann/Spinco SW 60 rotor. Sucrose solutions contained KOH at 0,2 M; Tris·HCl, pH 7.5, at 10 mM; NaCl at 0.5 M; EDTA at 1 mM; Sarcosyl NL 97 (Ciba-Geigy) at 0.075%.

RESULTS

DNA synthesis and ligation. The template-directed incorporation of s^4T_d into bacteriophage fd RF DNA was examined. DNA synthesis was catalyzed by E.coli DNA polymerase I. Priming oligonucleotides originated from a DNase digest of fd RF DNA. The conditions for DNA synthesis were almost exactly those used for the 'repair' synthesis of unmodified RF molecules (15). T_dTP was replaced by s^4T_dTP. T4 DNA ligase (or the E.coli enzyme) was present to allow conversion of nicked circular RF II DNA into covalently closed circular RF I DNA. (For details see Materials and Methods.) The reaction was monitored by incorporation of $[^3H]C_dMP$ or $[^{32}P]A_dMP$ into acid-precipitable material.

Fig. 1 shows that in the presence of s^4T_dTP synthesis started immediately after addition of DNA polymerase, but proceeded at a much lower rate than in the presence of T_dTP. Nevertheless the same amount of DNA was obtained in both cases. The reactions reached plateau values after about 1.5 h and 15 h, respectively. In a control assay, when neither T_dTP nor s^4T_dTP were present, no DNA synthesis was observed (fig. 1).

The intermediates and final products of the reaction were analyzed by agarose gel electrophoresis. Fig. 2a demonstrates that DNA synthesis in the presence of s^4T_dTP was initiated on all template molecules and led to covalently closed RF I DNA. The formation of this species was independently verified by a cellulosenitrate filter assay which makes use of the selective denaturation of RF II and RF III DNAs (32,33), and by velocitiy sedimentation in an alkaline sucrose gradient as shown in fig. 3. Minor amounts of linear full-length double-stranded RF III DNA were found (fig. 2a). As the template contained only traces of linear DNA it seems that the appearance of this

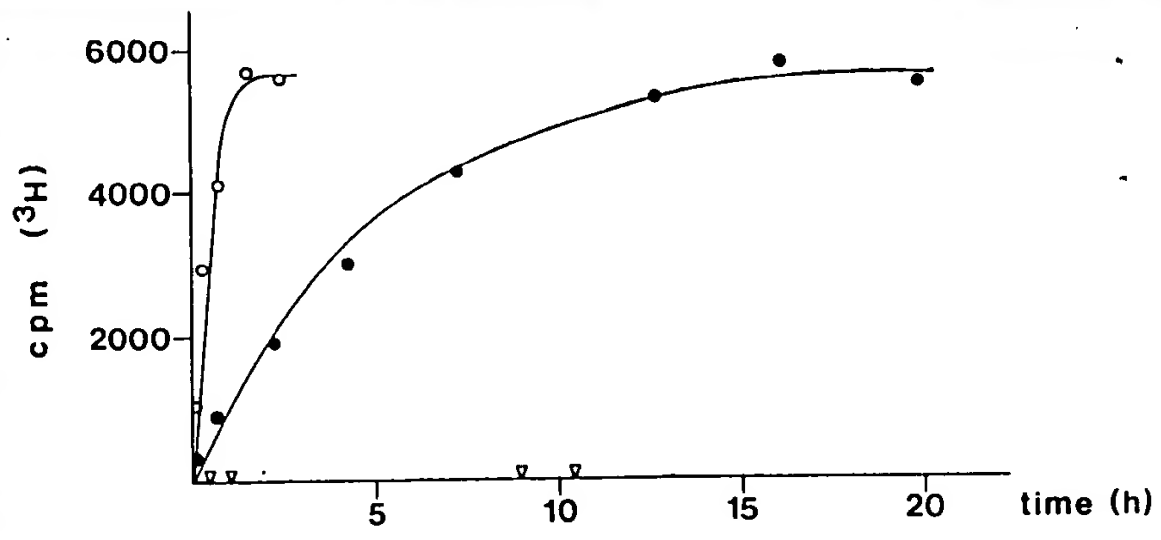


Figure 1: Time course of DNA synthesis catalyzed by DNA polymerase I. At times 5µl-aliquots were withdrawn from the reaction mixture and assayed for the incorporation of ^3H from $[^3H]C_dTP$ into acid-precipitable material. The reaction was carried out in the presence of A_dTP, G_dTP, C_dTP, and T_dTP (open circles); A_dTP, G_dTP, and C_dTP (triangles); A_dTP, G_dTP, C_dTP, and s^4T_dTP (full circles).

species was due to traces of endonuclease contaminations.

Agarose gel electrophoresis shows that only relatively small amounts of nicked circular RF II DNA (being precursors of RF I DNA) were present at any time samples were withdrawn from the reaction mixture. When normal DNA was synthesized under identical conditions considerably more RF II species were observed (fig. 2b). But when the rate of DNA synthesis was decreased by reducing the amount of DNA polymerase I about 3.5-fold, the conversion of RF II species into RF I molecules could be enhanced to about the same amount as observed in the presence of s^4T_dTP (fig. 2a).

When RF I DNA is synthesized under the conditions described above, a Gauss distribution of molecules with different topological winding numbers is obtained which can be separated by agarose gel electrophoresis in the absence of ethidium bromide (34,35). No significant difference in the electrophoretic pattern of normal and s^4T_d-substituted RF I DNAs was observed (fig. 2c).

During the isolation of RF I DNA we noticed that considerable amounts of this species were lost by conversion to

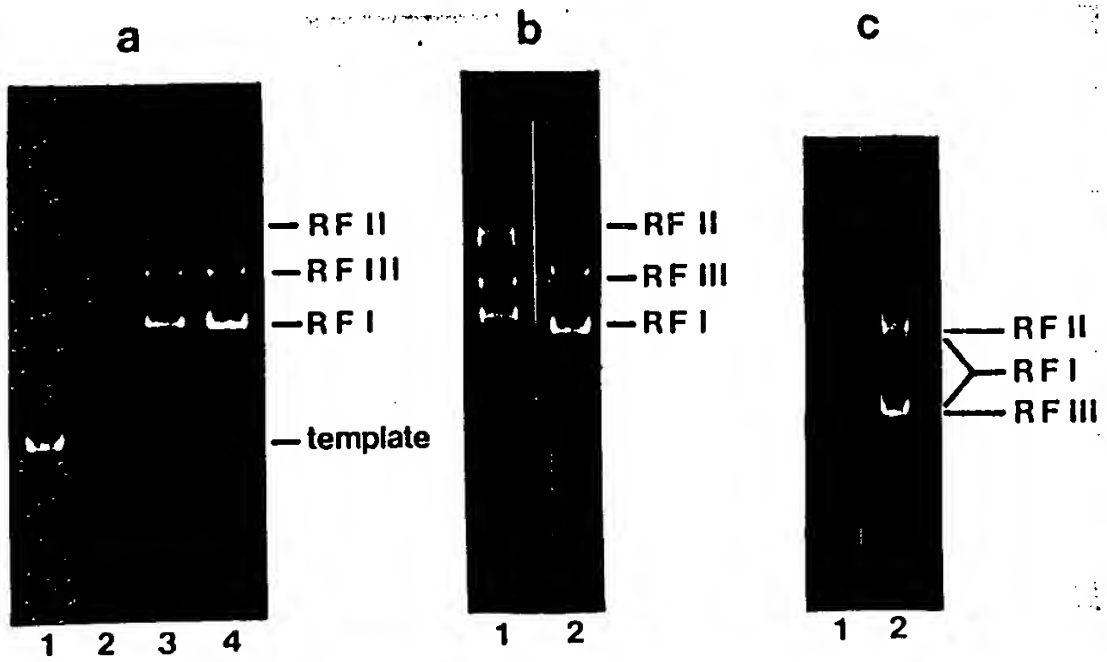


Figure 2: Analysis of the products of DNA synthesis by agarose gel electrophoresis. Migration is from top to bottom. a: Synthesis in the presence of s^4T_dTP. 3µl-samples were withdrawn at zero time (lane 1), 1 h 35 min (lane 2), 5 h 35 min (lane 3), 15 h 45 min (lane 4) and electrophoresed in buffer A containing ethidium bromide. b: Synthesis in the presence of T_dTP. 2 h after the amount of acid-precipitable radioactivity had reached a plateau the products were analyzed as above. DNA synthesis was catalyzed by 60 units/ml (lane 1) or 18 units/ml (lane 2) of DNA polymerase I. c: End products electrophoresed in the absence of ethidium bromide (buffer B). s^4T_d-substituted DNA (lane 1) and normal DNA (lane 2).

RF II. Further examination of this phenomenon revealed that nicking occured almost exclusively during ethanol precipitation and/or resolution of the pellet with a vortex mixer. This seems to indicate an increased susceptibility of the modified DNA to mechanical forces.

As the T_d-analogue displays an absorption maximum at 335 nm, the s^4T_d-content of the isolated RF I DNA was examined by UV spectroscopy. The absorption spectrum clearly indicated the presence of the modified base (fig. 4). In order to quantitate the amount of s^4T_dMP present in the DNA preparation the molecules had to be degraded monomers, as the percentage of hyperchromicity of this DNA was unknown. This was done at pH 8.0 by the subsequent actions of DNAse I and snake venom phosphodi-

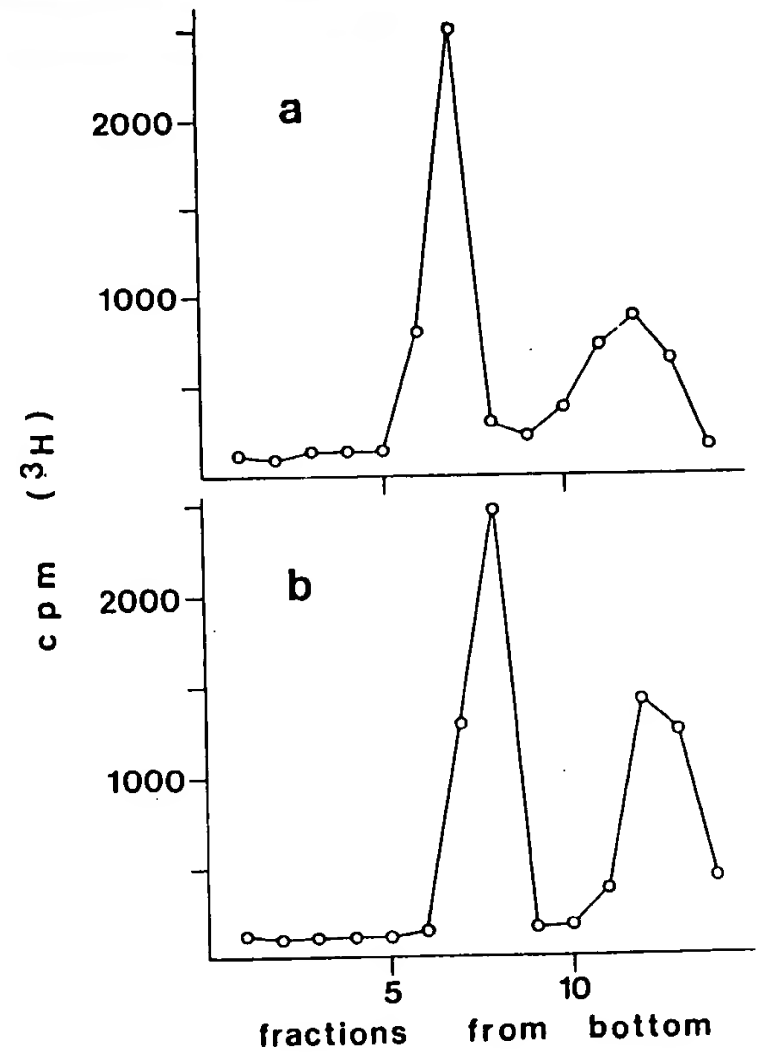


Figure 3: Analysis of the products of DNA synthesis by velocity sedimentation in an alkaline sucrose gradient. The fractions were assayed for acid-precipitable radioactivity. a: DNA synthesized in the presence of s^4T_dTP. b: DNA synthesized in the presence of T_dTP.

esterase. This treatment led to a shift of the short wavelength maximum from 255 nm to 259 nm. No significant shift of the long wavelength maximum was observed. An increase in absorbance of 62% and 115% at 260 nm and 335 nm, respectively, was found. This yields an absorbance ratio A_{260}/A_{335} of 4.35. The theoretical value was calculated using the following data for the molar absorption coefficients of the nucleotides: a) at 260 nm: A_d: 15400, G_d: 12010, C_d: 7050, T_d: 8400, and s^4T_d: 2500 (19);

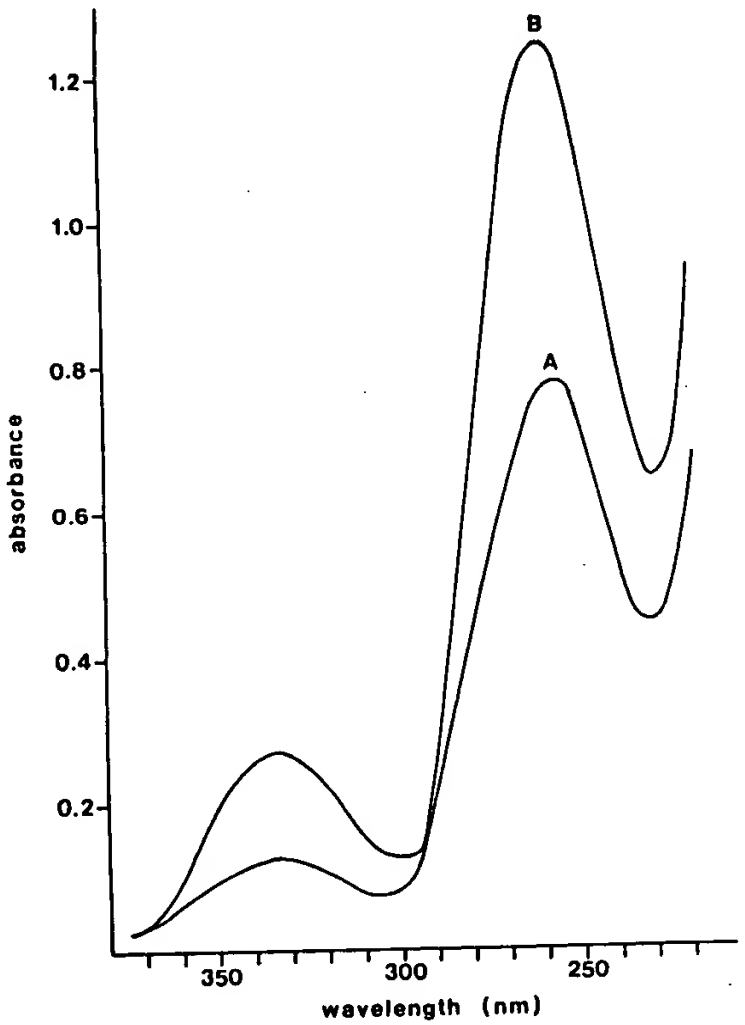


Figure 4: Ultraviolet absorption spectrum and hyperchromicity of fd RF I DNA synthesized in the presence of s^4T_dTP. The DNA was dissolved in 100 mM Tris·HCl, pH 8.0; 10 mM $MgCl_2$; 0,2 mM EDTA. A, before, and B, after degradation to mononucleotides.

b) at 335 nm: zero for the normal nucleotides and 21000 for s^4T_d (19). The values for the base composition of fd RF DNA were derived from the known sequence of the plus-strand: 34.5% A_d, 24.6% T_d, 20.7% C_d, 20.2% G_d. This leads to a theoretical absorbance ratio $A_{260}/A_{335} = 3.95$. It follows that a value of 4.53 corresponds to an 88% s^4T_dMP-substitution of the minus-strand.

To independently prove the s^4T_dMP content of the molecules DNA synthesis in the presence of $[\alpha\text{-}^{32}P]A_dTP$ was performed. The product was degraded by the subsequent actions of micrococcal

nuclease and spleen phosphodiesterase to yield labeled 3'-nucleotides. These were chromatographed on a silica gel thin layer. The digest of unmodified DNA was run as a control. With ethanol-1 M ammonium acetate, pH 7.5, 7:3 (v/v) as solvent s^4T_dMP and T_dMP were separated from each other and from the other nucleotides. As shown in fig. 5, DNA synthesized in the presence of s^4T_dTP in fact contained a new component which was not present in the normal DNA and had the chromatographic properties of s^4T_dMP (unlabeled 5'-s^4T_dMP run as additional control, not shown). Minor amounts of T_dMP, presumbly resulting from hydrolysis of the 4-keto-group, were detected. They figured up to about 14% as quantified by Cerenkov counting and densitometric evaluation. This is in good agreement with the spectroscopic analysis.

Restriction endonuclease cleavage. s^4T_d-containing fd RF DNA

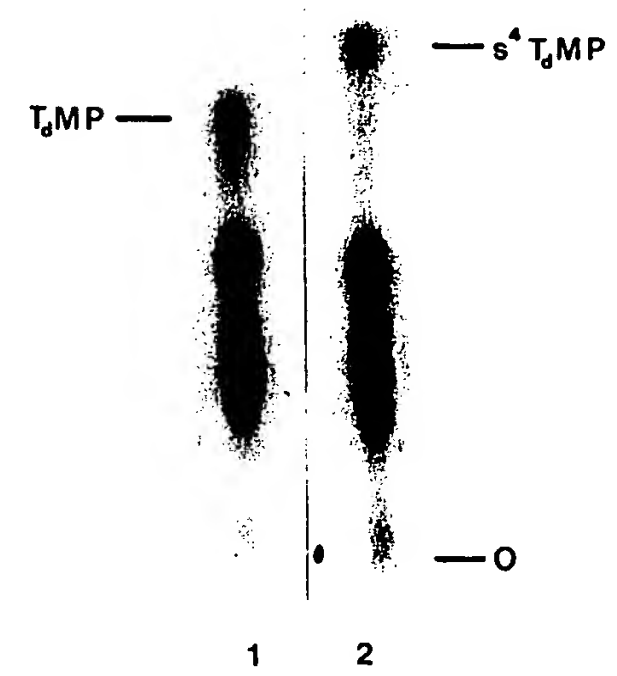


Figure 5: Chromatographic analysis of mononucleotides after hydrolysis of DNA. DNA was synthesized in the presence of $[\alpha\text{-}^{32}P]A_dTP$. It was digested subsequently by nuclease from St. aureus and spleen phosphodiesterase to yield 3'-N_dMPs. The products were chromatographed on silica gel thin layer. The appropriate 5'-N_dMPs were run as markers (not shown). Lanes 1 and 2 show DNA synthesized in the presence of T_dTP or s^4T_dTP, respectively. O=origin.

was incubated with restriction endonucleases Hpa II and Taq I, normally recognizing the sequences $(C/CGG)_d$ and $(T/CGA)_d$, respectively. The products were analyzed on 3.5% polyacrylamide/7M urea gels. As shown in fig. 6a, the modified DNA was cleaved not only by Hpa II which has only C_d and G_d in its recognition site, but also by Taq I normally cutting next to T_d. It was observed that the rates of cleavage by Hpa II were reduced by the DNA modification. The amount of enzyme had to be enhanced about 5-fold to achieve complete cleavage within the time neccessary to fragment unmodified DNA. At different Hpa II sites the cleavages rates were different. This effect was also observed for the restriction of unmodified fd RF DNA by Hpa II (our unpublished results).

s^4T_d substitution led to significantly altered electrophoretic mobilities of certain DNA fragments in polyacrylamide gels (fig. 6). The modified Hpa II-fragments D and E (0.652 and 0.648 kb) comigrated under the conditions applied, whereas the unsubstituted fragments were clearly separated. The same holds for Taq I-fragments G and H (presumably 0.381 and 0.357 kb). On

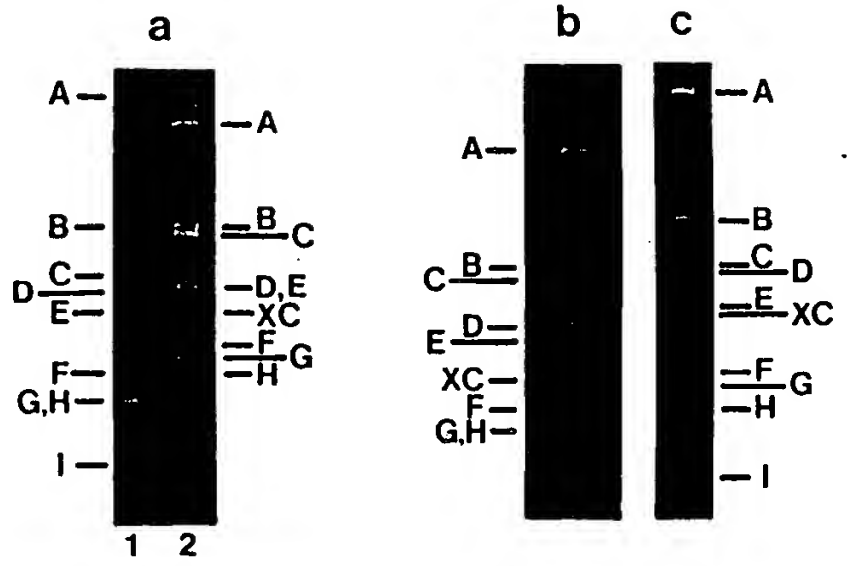


Figure 6: Cleavage of fd RF DNA by restriction endonucleases. The products were analyzed by electrophoresis in a 3.5% polyacrylamide gel. Migration is from top to bottom. XC = xylene cyanol. a: s^4T_d-substituted RF DNA cleaved by Taq I (lane 1) or Hpa II (lane 2). b: Normal RF DNA cleaved by Hpa II. c: Normal RF DNA cleaved by Taq I.

the other hand s^4T_d incorporation improved the separation of the Hpa II-fragments G and H (0.454 and 0.381 kb). These effects did not correlate with the s^4T_d-content of the fragments, but obviously were sequence specific.

DISCUSSION

Several years ago s^4T_dTP had first been synthesized and tested as substrate for DNA polymerases (18-20). It was found that templates of strictly alternating sequences like poly($A_d \cdot T_d$) and poly($A_d \cdot C_d$) were able to direct incorporation of s^4T_d into the complementary strand. Using the homopolymer template poly(A_d), however, Lezius reported inhibition of DNA synthesis in the presence of a 20-750fold excess of T_dTP over s^4T_dTP. This effect was explained by the assumption that incorporation of s^4T_d inhibits further primer elongation by DNA polymerase I. With activated calf thymus DNA only little incorporation of s^4T_d was observed when T_d or s^4T_d (no discrimination between these two possibilities was given) were the preceding nucleotides.

We re-examined the incorporation of s^4T_d into DNA replacing a merely defined substrate like calf thymus DNA by unique template molecules of known sequence. Using the circular single-stranded genome of bacteriophage fd to direct nucleotide incorporation, full-length complementary strand synthesis could not only be checked by chain-length determination but even more exactly by the formation of covalently closed circular RF I molecules in the presence of DNA ligase. We were able to show that in the presence of s^4T_dTP DNA synthesis by E.coli DNA polymerase I was considerably retarded, but not completely inhibited. It was initiated on all template molecules and led to the formation of full-length complementary strands in virtually all cases. The newly synthesized DNA was shown to be about 90% substituted by s^4T_d.

Inspection of the template sequence shows that it contains 5 $(A_d)_6$-, 12 $(A_d)_5$-, 33 $(A_d)_4$-, 79 $(A_d)_3$-, and 229 $(A_d)_2$-stretches. This indicates that s^4T_d-s^4T_d-sequences can in fact be synthesized by DNA polymerase I.

In our opinion two major reasons could account for the dis-

crepancies between our results and the earlier reports mentioned above. First, a homopolymer pair like poly(A_d)·poly(T_d) probably is structurally different from DNA. Hence, results obtained with this system might not be valid for DNA. Second, nuclease contamination of the DNA polymerase might be a critical point. It was found especially in restriction endonuclease reactions that the modified DNA was extremly susceptible to contaminating exonucleolytic activities. Furthermore we formerly observed that U_d-containing DNA was so rapidly degraded by exonuclease contaminations of DNA polymerase I preparations that we were not able to obtain RF I molecules in the presence of DNA ligase.

It seems likely that at least under certain conditions the conformation of s^4T_d-containing DNA differs from that of normal DNA. Although the $A_d \cdot s^4T_d$ base pair seems to be of the Watson-Crick type (36), replacement of the NH···O bond (2.9 Å) of the $A_d \cdot T_d$ pair by a NH···S bond (3.3 Å) (37) should give rise to a distortion of the modified base pair. In fact X-ray diffraction studies on poly($A_d \cdot s^4T_d$) fibers revealed a deviation of the $A_d \cdot s^4T_d$ pair from planarity (Saenger, W., personal communication). Furthermore it was noticed that s^4T_d-substituted DNA fragments show altered electrophoretic mobilities. As these effects were not proportional to the s^4T_d-content of the respective fragments, they cannot be explained by an altered net charge alone, but seem to reflect sequence-specific structural changes. Similar observations were also made for other base analogues (15). The enhanced susceptibility of the modified RF I DNA to mechanical forces might indicate an unusual conformational strain.

Another interesting difference between normal and s^4T_d-containing DNA is the magnitude of the hyperchromic effect. At 260 nm the hyperchromicity of the modified DNA totaled to about 62%. For normal DNA of similar base composition values around 80% were observed at a similar pH (38, and our own observations). This indicates that the incorporation of s^4T_d leads to a significant reduction of the stacking interactions of the normal nucleotides. This may be explained by a partially reduced overlap of p-orbitals of the heterocyclic bases due to the deviation of the $A_d \cdot s^4T_d$ base pair from planarity. The hyperchromicity at 335 nm indicates that the nucleotide analogue is involved in base stacking. The value of 115% is relatively large. It is rather small, however, compared to the respective value of 173% observed for poly($A_d \cdot s^4T_d$) (19). Interestingly the hyperchromic effect of this alternating polymer was also larger at 260 nm (80%).

The rate of primer elongation by DNA polymerase I in the presence of s^4T_dTP was significantly reduced. It is not known if this effect reflects the differences between T_d and its analogue in the interaction between the enzyme and the N_dTP, the enzyme and the base pair, or/and the enzyme and the primer terminus. There is some evidence that at least the third alternative might play a role. In the presence of DNA ligase s^4T_d-containing DNA is sealed faster than normal DNA. The ligation of the normal substrate can be improved by lowering the polymerase concentration. This seems reasonable as both enzymes compete for nicks (ligation vs. nick-translation). Obviously the ratio of binding of polymerase and ligase is lowered for the modified DNA. It follows that the affinity for nicks of at least one of the two proteins is changed with the s^4T_d-substituted substrate.

Our studies on fragmentaiton of the modified RF I DNA show that not only a restriction endonuclease recognizing sequences without T_d, like Hpa II, is able to specifically cleave this substrate, but that also an enzyme like Taq I, which normally cuts next to T_d, generates the fragments expected. So in principle the advantages of site-specific DNA cleavage can also be applied to s^4T_d-containing DNA.

The findings that the modified DNA is fragmented considerably more slowly by Hpa II, and that individual Hpa II-sites on fd RF DNA are cleaved at different rates, confirm our earlier observations that sequences outside the Hpa II-recognition site display a remarkable influence on the enzymatic activity (15).

Based on the results presented here it may be expected that DNA sequences in general are accessible to modification by s^4T_d. It should also be possible to site-specifically fragment the products. They may be used in studies on DNA-DNA and DNA-protein interaction. As it was shown that the analogue can be attacked by several reagents under rather mild conditions (e.g.,

39-41), they should also be versatile probes or precursors for further DNA modifications.

ACKNOWLEDGMENT

The authors wish to thank R. Frank for T4 ligase, H. Müller for DNA polymerase I, H. Mayer for Taq I, D. Müller for oligonucleotide primers, N. Hoppe for computer programming, and the Deutsche Forschungsgemeinschaft for financial support.

REFERENCES

1) RF I, II, and III DNA, duplex closed circular, nicked circular and linear 'replicative form' DNA species; DTE, dithioerythritol; ME, 2-mercaptoethanol. Nucleotide abbreviations follow IUPAC-IUB recommendations, Eur. J. Biochem. (1970) 15, 203-208, including s^4T_d for deoxy-4-thiothymidine.
2) Inman, R.B., and Baldwin, R.L. (1962) J. Mol. Biol. 5, 172 - 184
3) Chamberlin, M., Baldwin, R.L., and Berg, P. (1963), J. Mol. Biol. 7, 334 - 349
4) Cullen, B.R., and Bick, M.D. (1976) Nucleic Acids Res. 3, 49 - 62
5) Sagi, J., Brahms, S., Brahms, J., and Ötvös, L. (1979) Nucleic Acids Res. 6, 2839 - 2848
6) Lin, S., Lin, D., and Riggs, A.D. (1976) Nucleic Acids Res. 3, 2183 - 2191
7) Goeddel, D.V., Yansura, D.G., and Caruthers, M.H. (1978) Proc. Natl. Acad. Sci. USA 75, 3578 - 3582
8) Stahl, S.J., and Chamberlin, M.J. (1978) J. Biol. Chem. 253, 4951 - 4959
9) Marchionni, M.A., and Roufa, D.J. (1978) J. Biol. Chem. 253, 9075 - 9081
10) Hofer, B., and Köster, H. (1979) FEBS-Letters 102, 87 - 90
11) Simpson, R.B. (1979) Proc. Natl. Acad. Sci. USA 76, 3233 - 3237
12) Kallos, J., Fasy, T.M., Hollander, V.P., and Bick, M.D. (1979) FEBS-Letters 98, 347 - 350
13) Berkner, K.L., and Folk, W.R. (1979) J. Biol. Chem. 254, 2551 - 2560
14) Fasy, T.M., Cullen, B.R., Lutz, D., and Bick, M.D. (1980) J. Biol. Chem. 255, 1380 - 1387
15) Hofer, B., and Köster, H. (1980) Nucleic Acids Res. 8, in press
16) Müller, W., Weber, H., Meyer, F., and Weissmann, C. (1978) J. Mol. Biol. 124, 343 - 358
17) Flavell, R.A., Sabo, D.L.D., Bandle, E.F., and Weissmann, C. (1974) J. Mol. Biol. 89, 255 - 272
18) Lezius, A.G., and Scheit, K.H. (1967) Eur. J. Biochem. 3, 85 - 94
19) Lezius, A.G., and Rath, U. (1971) Eur. J. Biochem. 24, 163 - 167

tile probes or precursors for

rank for T4 ligase, H. Müller
Taq I, D. Müller for oligo-
computer programming, and the
or financial support.

closed circular, nicked cir-
form' DNA species; DTE,
ptoethanol. Nucleotide ab-
recommendations, Eur. J.
including s^4T_d for deoxy-4-

. (1962) J. Mol. Biol. 5,
, and Berg, P. (1963), J.
1976) Nucleic Acids Res. 3,
J., and Ötvös, L. (1979)
848
D. (1976) Nucleic Acids
and Caruthers, M.H. (1978)
3578 - 3582
J. (1978) J. Biol. Chem.
J. (1978) J. Biol. Chem.
9) FEBS-Letters 102, 87 - 90
l. Acad. Sci. USA 76,
der, V.P., and Bick, M.D.
350
(1979) J. Biol. Chem. 254,
D., and Bick, M.D. (1980)
) Nucleic Acids Res. 8,
., and Weissmann, C.
358
ndle, E.F., and Weissmann,
- 272
(1967) Eur. J. Biochem. 3,
1) Eur. J. Biochem. 24,

20) Lezius, A.G. (1971) Hoppe Seyler's Z. Physiol. Chem. 352, 491 - 500
21) Beck,E., Sommer, R., Auerswald, E.A., Kurz, C., Zink, B., Osterburg, G., Schaller, H., Sugimoto, K., Sugisaki, H., Okamoto, T., and Takanami, M. (1978) Nucleic Acids Res. 5, 4495 - 4503
22) Englund, P.T. (1971) in Procedures in Nucleic Acids Research, (Cantoni, G.L., and Davies, D.R., eds.), Vol. 2, pp. 864-874, Harper and Row, New York
23) Blöcker, H., Frank, R., and Köster, H. (1978) Liebigs Ann. Chem. 1978, 991 - 1006
24) Olivera, B.M., and Lehmann, I.R. (1967) Proc. Natl. Acad. Sci. USA 57, 1426 - 1433
25) Zimmermann, S.B., and Oshinsky, C.K. (1969) J. Biol. Chem. 244, 4689 - 4695
26) Modrich, P., and Lehmann, I.R. (1970) J. Biol. Chem. 245, 3626 - 3631
27) Scheit, K.H. (1968) Chem. Ber. 101, 1141 - 1147
28) Müller, D., Hofer, B., and Köster, H., submitted for publication
29) Josse, J., Kaiser, A.D., and Kornberg, A. (1961) J. Biol. Chem. 236, 864 - 875
30) Laskowski, Sr., M. (1971) in The Enzymes (Boyer, P.D. ed.), Vol. 4, pp. 313 - 328, Academic Press, New York - London
31) Wheeler, F.C., Fishel, R.A., and Warner, R.C. (1977) Anal. Biochem. 78, 260 - 275
32) Jansz, H.C., Pouwels, P.H., and Schiphorst, J. (1966) Biochem. Biophys. Acta 123, 626 - 627
33) Gray, C.P., Sommer, R., Polke, C., Beck, E., and Schaller, H. (1978) Proc. Natl. Acad. Sci. USA 75, 50 - 53
34) Depew, R.E., and Wang, J.C. (1975) Proc. Natl. Acad. Sci. USA 72, 4275 - 4279
35) Pulleyblank, D.E., Shure, M., Tang, D., Vinograd, J., and Vosberg, H.-P. (1975) Proc. Natl. Acad. Sci. USA 72, 4280 - 4284
36) Zipper, P. (1973) Eur. J. Biochem. 39, 493 - 498
37) Donohue, J. (1969) J. Mol. Biol. 45, 231 - 235
38) Mahler, H.R., Kline, B., and Mehrotra, B.D. (1964) J. Mol. Biol. 9, 801 - 811
39) Ziff, E.B., and Fresco, J.R. (1969) Biochemistry 8, 3242 - 3248
40) Favre, A., Yaniv, M., and Michelson, A.M. (1969) Biochem. Biophys. Res. Commun. 37, 266 - 271
41) Jida, S., Chung, K.C., and Hayatsu, H. (1973) Biochem. Biophys. Acta 308, 198 - 204

Pergamon

Bioorganic & Medicinal Chemistry Letters 10 (2000) 907–910

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

Incorporation of 4-Thiothymidine Into DNA by the Klenow Fragment and HIV-1 Reverse Transcriptase

Tata Venkata S. Rao, Martin T. Haber, Jane M. Sayer* and Donald M. Jerina

Laboratory of Bioorganic Chemistry, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA

Received 25 October 1999; accepted 16 February 2000

Abstract—The 5′-triphosphate of 4-thiothymidine (4S–TTP) is an excellent substrate for the Klenow fragment of *Escherichia coli* DNA polymerase I and HIV-1 reverse transcriptase with values of k_{cat}/K_m within a factor of ~3 of those for TTP. A large UV change ($\Delta\epsilon = -9770\ M^{-1}\ cm^{-1}$ at 340 nm) associated with incorporation of 4S–TMP into nucleic acid duplexes makes possible a rapid, continuous spectrophotometric assay of the reaction progress. © 2000 Published by Elsevier Science Ltd.

The kinetics of nucleotide incorporation into DNA duplexes have been extensively studied in attempts to understand the mechanism of action of DNA polymerases.[1] Recently, the role of hydrogen bonding and base pairing in the fidelity of DNA polymerization has been debated.[2] A nucleotide containing a difluorotoluene shape analogue of thymine, which lacks the ability to form any of the hydrogen bonds found normally in DNA base pairs, is enzymatically incorporated into DNA with good kinetics and high fidelity.[3]

4-Thiothymidine triphosphate[4] (4S–TTP) is incorporated into DNA by *Escherichia Coli* DNA polymerase I,[5] and synthetic oligonucleotides containing 4S–T have been used as probes of protein–DNA contacts.[6] Substitution of one oxygen of TTP by sulfur to form 4S–TTP may result in modification of one of the two hydrogen bonds found in the product dA-T base pair (Fig. 1), since the 1.68 Å length of the C–S double bond in 4S–T is significantly greater than the 1.23 Å length of the corresponding C–O double bond of thymidine.[7] Incorporation of one or two 4S–T's into a short oligonucleotide duplex has little effect on the melting temperature or circular dichroism spectra, consistent with only minimal effects on duplex DNA structure.[4b] Raman spectroscopy of oligonucleotide duplexes containing 4S–T indicates that the strength of the hydrogen bonds to 4S–T is similar to that of normal Watson–Crick hydrogen bonds.[6b] There is other evidence supporting the existence of N–H···S hydrogen bonds[8] in general, but the exact orientation of the dA-(4S–T) base pair in a modified DNA duplex has not been determined.

Substitution of sulfur for oxygen shifts the UV absorbance maximum from 267 nm for TTP to 335 nm for 4S–TTP, well removed from the background absorbance of DNA. Furthermore, incorporation of 4S–TTP into a DNA template–primer duplex results in significant hypochromicity (Fig. 2), which forms the basis for a convenient, continuous non-radiochemical assay for DNA polymerization.

The observed molar absorbance decrease at 340 nm is $\Delta\epsilon = 9770\ M^{-1}\ cm^{-1}$.[10] Incubations with duplexes designed to incorporate 2, 3 and 4 residues[11] of 4S–T showed that $\Delta\epsilon$ increased linearly with the number of 4S–T residues incorporated with a slope of $9770 \pm 150\ M^{-1}\ cm^{-1}$. This linearity is consistent with our previously reported results with normal nucleotides.[12] However, the value of the hypochromicity ($\Delta\epsilon$) at 340 nm is much larger than the value of $\Delta\epsilon$ at 275 nm[12] for TTP, where the large background absorbance of the DNA template-primer also interferes with the assay. For 4S–T, the hypochromicity at 340 nm is much larger than its hypochromicity at 260 nm.[6c]

Under *processive* conditions,[1] multiple nucleotides are incorporated into the DNA homopolymer duplex $p(dA)_{40-60}{\cdot}p(dT)_{20}$. In order to determine the average number of nucleotides incorporated into the product, the template-primer duplex was incubated with 4S–TTP and Klenow fragment until no more UV change was observed. The double-stranded product was isolated by

*Corresponding author. Tel.: +1-301-496-9893; fax: +1-301-402-0008; e-mail: jmsayer@helix.nih.gov

0960-894X/00/$ - see front matter © 2000 Published by Elsevier Science Ltd.
PII: S0960-894X(00)00123-2

precipitation with ethanol at 0 °C to remove excess 4S–TTP. Comparison of the product absorbance at 260 and 335 nm allowed calculation of the number of 4S–T residues incorporated. The most notable difference between 4S–TTP and TTP was a low overall incorporation (only 4.7 ± 1.0 nucleotides) of 4S–TTP by intact Klenow fragment. In contrast, an average of approximately 35 thymidine residues are incorporated by intact Klenow fragment into this template-primer, as measured by filter binding assay of ^{3}H-TTP incorporation.[12] (Note that the 20-mer primer presumably binds randomly to template (average length 50 nucleotides), so that the overhang is random unless the primer can change its position on the template.) For 4S–TTP the relative rates of incorporation and exonucleolytic proofreading apparently become approximately equal after incorporation of only a few 4S–T residues. In contrast, the incorporation rate is always much faster than the exonuclease rate for the natural substrate TTP, so that thymidine is incorporated until the end of the homopolymer duplex (or slightly beyond due to slippage). The mechanism resulting in lower incorporation of 4S–T into the duplex homopolymers by intact Klenow fragment is unknown, and may involve a decrease in the incorporation rate or an increase in the exonuclease rate with multiple, contiguous 4S–T residues, or a combination of both effects. With exonuclease-free Klenow fragment, approximately 37 4S–TTP residues were incorporated into DNA after an overnight reaction, comparable to the result with intact Klenow fragment and TTP. In general, a processivity of 20–25[1e] is observed for Klenow fragment depending on the template-primer.

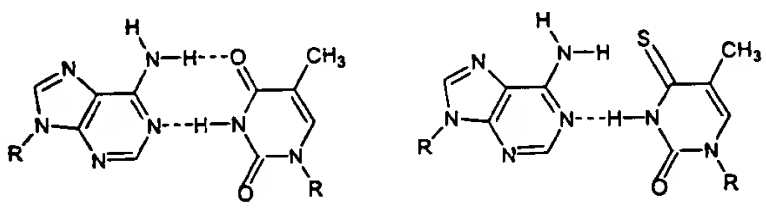


Figure 1. Base paired structures of dA-T and dA-(4S–T).

Kinetic parameters for the *processive* incorporation of 4S–TTP and TTP into the homopolymers are given in Table 1. These results are consistent with the reported k_{cat} for TTP of 3.8 s^{-1} obtained with the homopolymer, poly(dA)$_{1000}$(T)$_{10}$.[1d] Sulfur substitution decreases K_m by 3.6-fold while decreasing k_{cat} only slightly. The specificity constant k_{cat}/K_m shows that under these conditions 4S–TTP is a somewhat better substrate than TTP for Klenow fragment.

Results of the steady-state kinetics measurements under *non-processive* conditions[1] where a single nucleotide is incorporated into a DNA 9/20-mer template-primer by Klenow fragment or into a RNA/DNA template-primer by HIV-1 reverse transcriptase (RT) are given in Table 2. With Klenow fragment, both k_{cat} and K_m increase with 4S–TTP, resulting in no change in k_{cat}/K_m relative to that of TTP. The observed, small increase in K_m may reflect somewhat poorer binding of the altered substrate by the binary complex of enzyme and DNA duplex. The increase in k_{cat} suggests that the rate of dissociation of the oligonucleotide product from the enzyme is increased upon incorporation of 4S–T relative to T, since this dissociation step is known to be rate limiting for the non-processive reaction of the normal nucleoside triphosphates with this same duplex.[1c]

Table 1. Steady-state parameters for DNA polymerization by Klenow fragment under processive conditions with the homopolymer template-primer duplex, p(dA)$_{40-60}$·p(dT)$_{20}$[a]

Substrate	K_m (μM)	k_{cat} (s^{-1})	10^{-5} k_{cat}/K_m (M^{-1} s^{-1})
TTP	5.0	3.0	6.0
4S–TTP	1.4	2.3	15.7

[a]Kinetic measurements were conducted[13] with Klenow fragment using 1 μM p(dA)$_{40-60}$·p(dT)$_{20}$ as template primer. Assays using 0.2–15 μM ^{3}H-TTP (2.0 Ci/mmol) as substrate were done by radiochemical filter binding assay.[15] Assays using 0.5–50 μM 4S–TTP as substrate were conducted by following the change in absorbance at 340 nm in microspectrophotometric cells.

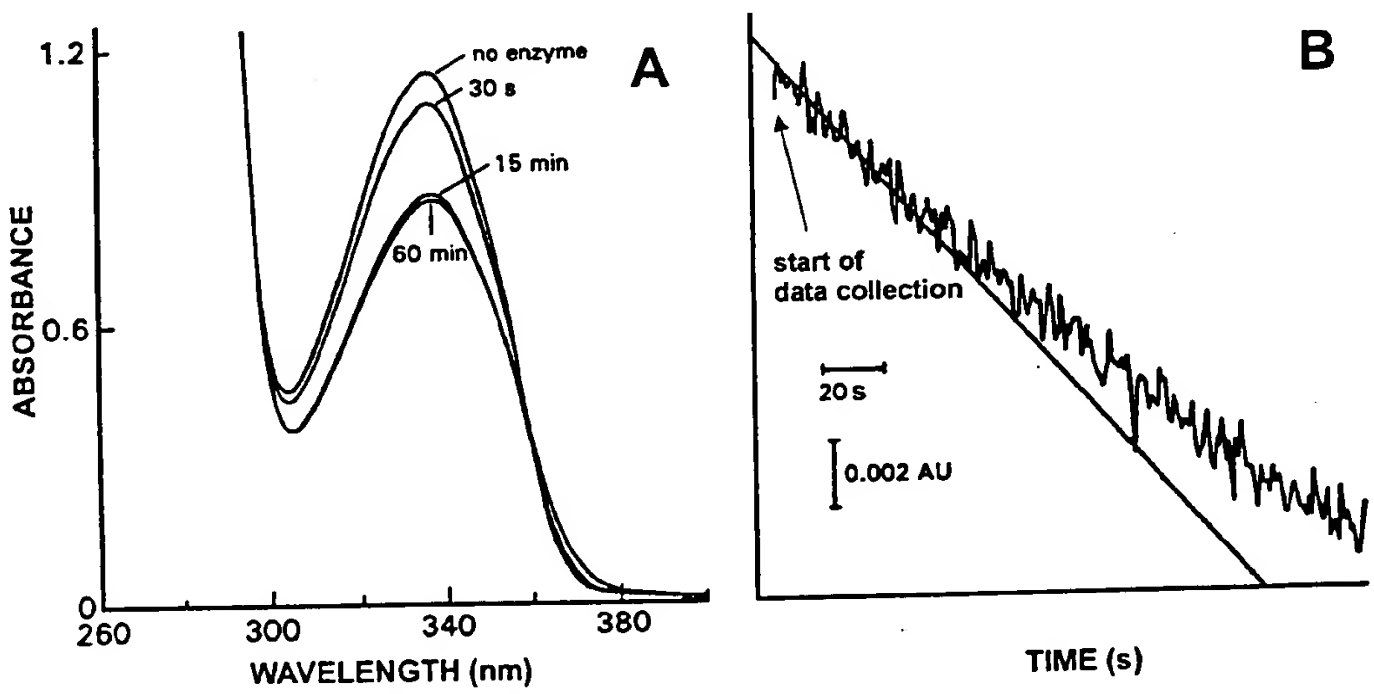


Figure 2. UV spectral change (A) and time course (B) for incorporation of one 4S–TTP into a 9/20-mer DNA duplex[9] by the Klenow fragment. Reaction conditions:[13] (A) 52 μM 4S–TTP, 25 μM duplex and 40 nM exonuclease-free Klenow fragment, and (B) 5 μM 4S–TTP, 10 μM duplex and 10 nM Klenow fragment.

Table 2. Steady-state parameters for DNA polymerization by Klenow fragment and HIV-1 reverse transcriptase (RT) under non-processive conditions with short template-primer duplexes

Enzyme	Substrate	K_m (μM)	k_{cat} (s^{-1})	$10^{-5}\ k_{cat}/K_m$ ($M^{-1}\ s^{-1}$)
Klenow fragment[a]	TTP	0.51	0.37	7.3
Klenow fragment	4S–TTP	1.82	1.73	9.5
HIV-1 RT[b]	TTP	0.36	1.02	28.0
HIV-1 RT	4S–TTP	0.50	0.44	8.8

[a]Spectrophotometric assays with 10 nM Klenow fragment were conducted[13] with 10 μM DNA template-primer[9] and 2.5–75 μM 4S–TTP. Radiochemical assays with 4 nM Klenow fragment were conducted under the same conditions with 10 μM DNA template-primer and 0.2–15 μM ^{3}H-TTP (7.2 Ci/mmol). The rate of thymidine incorporation was determined with the radiochemical filter binding assay.[15]
[b]Spectrophotometric assays with RT were carried out under conditions similar to those described[16] in 50 mM Tris buffer, pH 7.4, containing 50 mM NaCl and 10 mM $MgCl_2$ with 5 μM DNA-9-mer (primer)/RNA-20-mer (5′-AAACCCUUGGACGGCUGCGA-OH, template), 0.25–25 μM 4S–TTP and 14 nM HIV-1 RT. Radiochemical assays were conducted under the same conditions with 5 μM DNA-9-mer/RNA-20-mer, 0.1–20 μM ^{3}H-TTP (18 Ci/mmol) and 1.0 nM HIV-1 RT. HIV-1 RT (MW 110000 Da) consisting of equimolar amounts of subunits p66 and p51 was a gift from Dr. Samuel Wilson.

To investigate further the nature of the rate determining step for incorporation of 4S–TTP, pre-steady-state kinetics of the reaction were measured on a msec time scale.[14] With related DNA duplexes, nucleotide incorporation by the Klenow fragment exhibits a burst[1c,1d] with a rate of 50 s^{-1}, which is much faster than product release. This fast burst step is assigned to a conformational change in the enzyme that occurs after substrate binding and before phosphodiester bond formation, which is even faster.[1f] A second slow conformational change follows the bond-formation step before product release.[1g] The reaction of 4S–TTP showed no burst, and proceeded with an initial incorporation rate of about 2–3 s^{-1}, which is comparable to the rate of turnover in the steady state. A reasonable interpretation is that dissociation of the oligonucleotide product is not appreciably slower than the preceding step(s) in the catalytic process for the Klenow fragment with 4S–TTP as substrate. Thus with 4S–TTP as substrate, there is no evidence for rate limiting dissociation of the oligonucleotide product from the enzyme, and either the conformational changes or bond formation may be at least partially rate limiting.

With HIV-1 RT, comparison of 4S–TTP with TTP indicates little or no change in K_m and a small decrease in k_{cat} with a resultant decrease in k_{cat}/K_m (Table 2). The decrease in k_{cat} suggests a decrease in the dissociation rate for the modified substrate, since product release is rate limiting for non-processive kinetics of TTP with this polymerase also.[17] However, this kinetic scheme may not apply to the 4S–TTP substrate (see above).

In summary, the chromophoric substrate 4S–TTP is well suited for use in a rapid, continuous optical assay to replace slow and time intensive single-point radiochemical assays for polymerase-mediated replication reactions in many applications such as enzyme purification and high throughput screening. Oligonucleotide templates or nucleoside triphosphates containing the fluorescent analogue 2-aminopurine have also been used in continuous DNA polymerase assays.[18] We observed little change (<4-fold for K_m and ≤3-fold for k_{cat}/K_m) in the steady-state kinetic parameters for enzymatic incorporation of 4S–TTP relative to TTP into a DNA duplexes under both processive and non-processive conditions. This is expected since X-ray crystallographic studies of the polymerase tertiary complex indicate that the polymerase has little interaction with the major groove side of the DNA at the insertion site.[19] The total extent of incorporation of 4–ST into a homopolymer primer by intact Klenow fragment was much lower than that of T. This observation is presumably a result of a slowing of the polymerization relative to the exonuclease rate in the presence of the modified oligonucleotide, since similar extents of incorporation of 4–ST and T were observed with *exo*-Klenow fragment where no exonuclease activity is present. Pre-steady-state kinetics of incorporation of a single nucleotide residue into a synthetic template-primer by Klenow fragment under conditions that give a rapid burst of T incorporation[1c,1d] provided no evidence for such a burst with 4–ST. Thus, in contrast with the normal oligonucleotide, a step other than dissociation of the 4–ST modified oligonucleotide duplex from the enzyme may be rate determining.

Acknowledgements

We thank Dr. Richard N. Armstrong and Dr. James Sellers for the use of their stopped-flow spectrophotometers.

References and Notes

1. (a) Carroll, S. S.; Benkovic, S. J. *Chem. Rev.* **1990**, 1291. (b) Benkovic, S. J.; Cameron, C. E. *Methods Enzymol.* **1995**, *262*, 257. (c) Kuchta, R. D.; Mizrahi, V.; Benkovic, P. A.; Johnson, K. A.; Benkovic, S. J. *Biochemistry* **1987**, *26*, 8410. (d) Eger, B. T.; Kuchta, R. D.; Carroll, S. S.; Benkovic, P. A.; Dahlberg, M. E.; Joyce, C. M.; Benkovic, S. J. *Biochemistry* **1991**, *30*, 1441. (e) Eger, B. T.; Benkovic, S. J. *Biochemistry* **1992**, *31*, 9227. (f) Polesky, A. H.; Dahlberg, M. E.; Benkovic, S. J.; Grindley, N. D. F.; Joyce, C. M. *J. Biol. Chem.* **1992**, *267*, 8417. (g) Dahlberg, M. E.; Benkovic, S. J. *Biochemistry* **1991**, *30*, 4835.
2. Goodman, M. F. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 10493.
3. Moran, S.; Ren, R. X.-F.; Kool, E. T. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 10506.
4. (a) 4-Thiothymidine was prepared as described in ref 4b except that the hydroxyl groups were protected as the *tert*-butyldimethylsilyl derivatives rather than the acetates. After deprotection with *tert*-butylammonium fluoride, 4-thiothymidine was converted to the 5′-*O*-triphosphate as described in ref 4c. The product was purified by HPLC on an ion-exchange column (SynchroPak AX100, 21.2 × 250 mm) eluted with a two step gradient at a flow rate of 8 mL/min. The first step is from 0 to 60% B in 15 min followed by increasing B to 100% in 5 min (solvent A, 0.025 M Et_3N–H_2CO_3, pH 7.5; solvent B, 0.5 M Et_3N–H_2CO_3, pH 7.5). The HPLC peak eluting at 25 min was collected, and the aqueous solution was concentrated

at ambient temperature. Periodically, the solution pH was checked and maintained at 7.5. Methanol was added to the syrup and evaporation continued. The resulting syrup was dissolved in a minimum volume of 0.1 M Et_3N–H_2CO_3, pH 7.5, and chromatographed on a reverse phase Hamilton PRP-1 column (7.0 × 305 mm) by ramping the gradient from 100% A to 100% B in 30 min (solvent A: 0.1 M Et_3N–H_2CO_3; solvent B: 50:50 Et_3N–H_2CO_3:CH_3CN) in order to remove residual inorganic pyrophosphate. The peak eluting at 10 min was collected, and the solvent was carefully evaporated at ambient temperature. The triphosphate was characterized by ^{31}P NMR (D_2O, phosphoric acid external standard): δ −14.8, −2.95 and 2.2 ppm. Purity was confirmed by TLC as described in ref 4c. Concentrations of 4S–TTP were determined spectrophotometrically at 335 nm, $\epsilon = 22\,300\ M^{-1}\ cm^{-1}$ (ref 4d). (b) Connolly, B. A.; Newman, P. C. *Nucleic Acids Res.* **1989**, *17*, 4957. (c) Kovacs, T.; Otvos, L. *Tetrahedron Lett.* **1988**, *29*, 4525. (d) Fox, J. J.; Von Praag, D.; Wempen, I.; Doerr, I .L.; Cheong, L.; Knoll, J. E.; Eidinoff, M. L.; Bendich, A.; Brown, G. B. *J. Am. Chem. Soc.* **1959**, *81*, 178.

5. (a) Hofer, B.; Koster, H. *Nucleic Acids Res.* **1981**, *9*, 753. (b) Lezius, A. G.; Scheit, K. H. *Eur. J. Biochem.* **1967**, *3*, 85.

6. (a) Bartholomew, B.; Braun, B. R.; Kassavetis, G. A.; Geiduschek, E. P. *J. Biol. Chem.*, **1994**, *269*, 18090. (b) Thorogood, H.; Waters, T. R.; Parker, A. W.; Wharton, C. W.; Connolly, B. A. *Biochemistry* **1996**, *35*, 8723. (c) Newman, P. C.; Nwosu, V. U.; Williams, D. M.; Cosstick, R.; Seela, F.; Connolly, B. A. *Biochemistry* **1990**, *29*, 9891.

7. Szabo, T.; Noreus, D.; Norrestam, R.; Stawinski, J. *Nucleic Acids Res.* **1993**, *21*, 3921.

8. Donohue, J. *J. Mol. Biol.* **1969**, *45*, 231.

9. Equimolar amounts of the template strand DNA 20-mer (5′-AAACCCTTGGACGGCTGCGA-OH) and primer strand DNA 9-mer (5′-TCGCAGCCG-OH) were mixed and annealed as described in ref 1c. Oligonucleotide concentrations were estimated based on $A_{260\ nm}$ using the program Oligo (National Biosciences, Plymouth, MN).

10. The value of Δε was obtained by incubating 25 μM template-primer (DNA-9-mer/DNA-20-mer) and excess 4S–TTP with exonuclease-free Klenow fragment (Unites States Biochemical) under standard buffer conditions (ref. 13) and allowing the reaction to go to completion. Exonuclease-free Klenow fragment was used since with intact Klenow fragment the total absorbance change was smaller and changes in absorbance were observed to reverse direction with prolonged reaction times, indicating slow hydrolysis of the product.

11. The following three template DNA strands were used: 5′-AAACCCTTGAACGGCTGCGA-OH, 5′-AAACCCTTAAACGGCTGCGA-OH, and 5′-AAACCCTAAAACGGCTGCGA-OH, which when annealed with primer 9-mer in the final template-primer duplex contain sites for incorporation of 2, 3 and 4 (thio)thymidines, respectively, opposite to the underlined A's.

12. Baillon, J. G.; Nashed, N. T.; Sayer, J. M.; Jerina, D. M. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 1014.

13. Reactions were conducted in 50 mM Na MOPS buffer (pH 7.4), 50 mM KCl, 10 mM $MgCl_2$, 1 mM DTT and 0.5 mM EGTA, as used for previous kinetic experiments with Klenow fragment and similar template-primer duplexes (ref 1c). Kinetic parameters (25 °C) were obtained by fitting the Michaelis–Menten equation to initial rates measured at several triphosphate concentrations using the program Enzfitter (Biosoft, Cambridge, UK).

14. (a) Stopped-flow spectrometers from Kin-Tek Instruments (ref 14b), with two identical loading syringes (actuated by either nitrogen gas pressure or a stepping motor) leading to a rapid (~1.6 ms) mixing chamber connected to a stop syringe, were used. The stop chamber was irradiated by a variable wavelength monochromatic light source set to 340 nm and monitored by a photodiode connected to a computer. Syringe A contained 4 μM Klenow fragment and 20 μM DNA-9-mer/DNA-20-mer (ref 9) in MOPS buffer. Syringe B contained 5–60 μM 4S–TTP in MOPS buffer. Data were collected for time periods ranging from 0.1 s to up to 10 s. The results of several runs at each substrate concentration were averaged. (b) Johnson, K. A. *Methods Enzymol.* **1986**, *134*, 677.

15. Altman, S.; Lerman, L. S. *J. Biol. Chem.* **1970**, *50*, 235.

16. Kati, W. M.; Johnson, K. A.; Jerva, L. F.; Anderson, K. S. *J. Biol. Chem.* **1992**, *267*, 25988.

17. Reardon, J. E. *Biochemistry* **1992**, *31*, 4473.

18. (a) Bloom, L. B.; Otto, M. R.; Beechem, J. M.; Goodman, M. F. *Biochemistry* **1993**, *32*, 11247. (b) Frey, M. W.; Sowers, L. C.; Millar, D. P.; Benkovic, S. J. *Biochemistry* **1995**, *34*, 9185.

19. (a) Brautigam, C. A.; Steitz, T. A. *Curr. Opin. in Struct. Biol.* **1998**, *8*, 54. (b) Kunkel, T. A.; Wilson, S. H. *Nat. Struct. Biol.* **1998**, *5*, 95.

protein bound to pRb translation products ≥28K, but did not show increased affinity for translation products ≥56K, as seen with GST-E1A (Fig. 4*a*). Furthermore, no binding to GST-E1A and GSTΔ*myc* protein was observed with an *in vitro*-translated pRb polypeptide that contained a C terminus at amino acid 667 of the wild-type protein (data not shown). These data suggest that the binding sites for E1A and c-*myc* protein pRb may include similar regions in the C terminus of pRb.

To investigate further whether c-*myc* and DNA tumour virus-encoded proteins bind to similar sites on pRb, we did a competition experiment. We used human papilloma virus 16 (HPV-16) E7 protein for this experiment as this protein binds strongly to pRb[15]. Bacterially synthesized HPV-16 E7 protein or an HPV-16 E7 peptide that contained the pRb binding site on the E7 protein were incubated with *in vitro*-translated pRb. After this, GSTΔ*myc* protein or GST-E1A protein was added and binding of pRb to GST fusion proteins was assayed. Both HPV-16 E7 protein and HPV-16 E7 peptide, but not control peptide, competed efficiently for binding of both GSTΔ*myc* and GST-E1A protein to pRb (Fig. 4*b*). We conclude that c-*myc*, HPV-16 E7 and adenovirus E1A proteins have overlapping binding sites on pRb.

The C-terminal region of pRb to which c-*myc* binds is frequently deleted in human cancer[16–18]. This raised the possibility that in these tumour cells the interaction between c-*myc* protein and pRb is perturbed. To test this hypothesis, we used a cloned pRb cDNA from the J82 bladder carcinoma cell line which carries a 35-amino-acid C-terminal deletion in pRb (amino acids 697–731; ref. 16). *In vitro*-translated wild-type pRb and mutant J82 pRb was subsequently tested for binding to GSTΔ*myc*. The mutant pRb from the J82 bladder carcinoma binds both GST-E1A and GSTΔ*myc* with greatly reduced affinity (Fig. 4*c*). We conclude that the 35-amino-acid deletion of pRb in the J82 carcinoma cell line greatly alters the ability of pRb to bind to c-*myc* protein.

Interesting parallels exist between the c-*myc* and the adenovirus E1A genes. Both encode nuclear phosphoproteins that cooperate with an activated *ras* oncogene in transformation[19,20]. Additionally, *myc* and E1A are structurally distantly related[21,22]. We now extend these similarities by showing that c-*myc* protein, like E1A, can bind to pRb. Because full-length c-*myc* protein is highly insoluble, it was necessary to study the interaction between c-*myc* protein and pRb *in vitro* with a truncated soluble form of c-*myc* protein. Our findings are nevertheless intriguing because the region of pRb to which c-*myc* protein binds is frequently deleted in human tumours[16–18]. This suggests that c-*myc* protein interacts with a physiologically important site of pRb.

The ability of viral oncogenes to transform is intimately linked to their ability to bind pRb[23–24]. Our finding that the HPV-16 E7-encoded protein competes with c-*myc* protein for binding to pRb suggests that both proteins control cell proliferation by binding to overlapping sites on pRb. In this regard it is noteworthy that either antisense c-*myc* oligonucleotides or pRb overexpression results in a block in cell cycle progression from G1 to S phase (ref. 25, and S. Friend, personal communication). On the basis of our observations, it is tempting to speculate that dominant and recessive oncogenes can cooperate through direct binding to control progression through the cell cycle. □

Received 22 May; accepted 15 July 1991.

1. Lüscher, B. & Eisenman, R. N. *Genes Dev.* **4**, 2025–2035 (1990).
2. Murre, C. M. M., McCaw, P. S. & Baltimore, D. *Cell* **56**, 777–783 (1989).
3. Blackwood, E. M. & Eisenman, R. N. *Science* **251**, 1211–1217 (1991).
4. Blackwell, T. K. *et al. Science* **250**, 1149–1151 (1990).
5. Pendergast, G. C. & Ziff, E. B. *Science* **251**, 186–189 (1991).
6. Stone, J. *et al. Molec. cell. Biol.* **7**, 1697–1709 (1987).
7. Sugiyama, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 9144–9148 (1989).
8. Moore, J. P., Hancock, D. C., Littlewood, T. D. & Evan, G. I. *Oncogene Res.* **2**, 65–80 (1987).
9. Smith, D. B. & Johnson, K. S. *Gene* **67**, 31–40 (1988).
10. Kaelin, W. G. *et al. Cell* **64**, 521–532 (1991).
11. Lee, W. H. *et al. Nature* **329**, 642–645 (1987).
12. Whyte, P. *et al. Nature* **334**, 124–129 (1988).
13. Friend, S. H. *et al. Nature* **323**, 643–646 (1986).
14. Hu, Q., Dyson, N. & Harlow, E. *EMBO J.* **9**, 1147–1155 (1990).
15. Dyson, N., Howley, P. M., Münger, K. & Harlow, E. *Science* **243**, 934–936 (1989).
16. Horowitz, J. M. *et al. Science* **243**, 937–940 (1989).
17. Yandell, D. W. *et al. New Engl. J. Med.* **321**, 1689–1695 (1989).
18. Dunn, J. M. *et al. Molec. cell. Biol.* **9**, 4596–4604 (1989).
19. Land, H., Parada, L. F. & Weinberg, R. A. *Nature* **304**, 596–602 (1983).
20. Ruley, H. E. *Nature* **304**, 602–606 (1983).
21. Ralston, R. & Bishop, J. M. *Nature* **306**, 803–806 (1983).
22. Figge, J., Wester, T., Smith, T. F. & Paucha, E. *J. Virol.* **62**, 1814–1818 (1988).
23. Whyte, P., Williamson, N. M. & Harlow, E. *Cell* **56**, 67–75 (1989).
24. De Caprio, J. A. *et al. Cell* **54**, 275–283 (1988).
25. Heikkila, R. *et al. Nature* **328**, 445–449 (1987).
26. Rustgi, A. K., Van't Veer, L. J. & Bernards, R. *Proc. natn. Acad. Sci. U.S.A.* **87**, 8707–8710 (1990).
27. Windle, J. J. *et al. Nature* **343**, 665–669 (1990).
28. Jones, R. E. *et al. J. biol. Chem.* **265**, 12782–12785 (1990).

ACKNOWLEDGEMENTS. We thank M. Billaud, A. Fattaey, T. Frebourg, D. Hill, W. Kaelin, S. Shin, T. Tsai and L. van't Veer for GST fusion proteins, D. Patrick for HPV E7 protein, R. Kingston for c-*myc in vitro* expression vector, J. Stone for mutant c-*myc*, plasmids J. Horowitz for the pRb plasmid, E. Harlow for discussions and critical reading of the manuscript, and M. Dembski for expert technical assistance. This work was supported by grants form the National Institutes of Health, the Searle Scholarship Foundation (R.B.), and an Institutional Grant from the Massachusetts General Hospital (N.D.). A.K.R. was supported by the American Cancer Society and by a Clinical Investigator Award.

Preferential DNA secondary structure mutagenesis in the lagging strand of replication in *E. coli*

Thuan Q. Trinh* & Richard R. Sinden†

Department of Molecular Genetics, Biochemistry and Microbiology, University of Cincinnati College of Medicine, 231 Bethesda Avenue, Cincinnati, Ohio 45267-0524, USA

WHEN present in single-stranded DNA, palindromic or quasi-palindromic sequences have the potential to form complex secondary structures, including hairpins, which may facilitate interstrand misalignment of direct repeats and be responsible for diverse types of replication-based mutations, including deletions, additions, frameshifts and duplications[1–5]. In regions of palindromic symmetry, specific deletion events may involve the formation of a hairpin or other DNA secondary structures which can stabilize the misalignment of direct repeats[1,2]. One model suggests that these deletions occur during DNA replication by slippage of the template strand and misalignment with the progeny strand[6,7]. The concurrent DNA replication model, involving an asymmetric dimeric DNA polymerase III complex which replicates the leading and lagging strands[8], has significant implications for mutagenesis. The intermittent looping of the lagging strand template, and the fact that the lagging strand template may contain a region of single-stranded DNA the length of an Okazaki fragment, provides an opportunity for DNA secondary-structure formation and misalignment. Here we report our design of a palindromic fragment to create an 'asymmetric palindromic insert' in the chloramphenicol acetyltransferase gene of plasmid pBR325. The frequency with which the insert was deleted in *Escherichia coli* depends on the orientation of the gene in the plasmid. Our results suggest that replication-dependent deletion between direct repeats may occur preferentially in the lagging strand.

Plasmid pBR325, a ColE1-derived plasmid, has a unidirectional origin of replication and requires many host-encoded proteins for replication[9]. The transcribed strand of the chloramphenicol acetyltransferase (*CAT*) gene is the leading strand template strand and the non-coding strand the lagging strand. Reversing the direction of the *CAT* gene changes the relationship between the coding strand and the leading and lag-

† To whom correspondence should be addressed.
* Present address: Merck, Sharp, and Dohme Research Laboratories, Department of Biochemistry and Biophysics PO Box 2000, Rahway, New Jersey 07065, USA.

```
Plasmids    Nucleotide sequence at the EcoRI site.

                      <-----      <---- <-----
pBR325p1    gctcatccggAATTCCTATAGGAATTCTaattccgtatggca
                    -------><--------

                    <----- <----      <-----
pBR523p1    tgccatacggAATTAGAATTCCTATAGGaattccggatgagc
                    -------><--------

                      <-----           <-----
pBR325n1    gctcatccggAATTCGTCTGATGCACGaattccgtatggca

                      <-----           <-----
pBR523n1    tgccatacggAATTCGTGCATCAGACGaattccggatgagc
```

FIG. 1 Nucleotide sequence of the plasmids around the *Eco*RI site in the *CAT* gene of pBR325 and pBR523. Shown are the sequences of the 'deletion insert' (capital letters) and the surrounding *CAT* gene sequence (lower-case). Inserts were chemically synthesized as *Eco*RI fragments and cloned into pBR325. To construct the pBR523 derivatives, the *CAT* gene was reversed by cutting with *Asu*II and religating. A difference in the *Pvu*II restriction patterns of pBR325 and pBR523 derivatives was used to screen potential pBR523 clones. The nucleotide sequence of all plasmids was confirmed by DNA sequencing. The various direct repeats are indicated by arrows above the sequence pointing in the direction of replication of the leading strand. The inverted repeat in the p1 insert is indicated by arrows below the sequence.

template strands of replication. Reversing the gene, rather than reversing the insert within the *CAT* gene, preserves the relationship between the insert and flanking DNA sequence that might influence the frequency of mutation.

The frequency of spontaneous deletion between direct repeats represents the sum of all possible events that result from homologous or illegitimate recombination, DNA repair activities or errors of DNA replication. If deletion between direct repeats occurs with equal frequency in the leading and lagging strands, there would be no effect of reversing the *CAT* gene. During replication, if deletion occurs preferentially in either the leading or lagging strand, deletion might occur at different frequencies. For a 'symmetric deletion insert', the mechanism of deletion or the DNA secondary structures that may be involved in deletion will be the same in either the leading or lagging strand. Consequently, the frequency of deletion should be unchanged by reversing the *CAT* gene. But for an 'asymmetric deletion insert', the DNA secondary structures that may be involved in deletion will be different in the leading and lagging strands, and thus a different deletion frequency will be observed when the *CAT* gene is reversed.

We have designed several 17- and 18-base-pair symmetric and asymmetric deletion inserts and have shown that the frequency of deletion from the *Eco*RI site in the *CAT* gene in pBR325 can vary over more than a 100-fold range in *recA*$^-$ *E. coli*. Reversion to chloramphenicol resistance (Cm^r) occurs primarily by deletion between the direct repeats, which include the *Eco*RI sites, resulting in a restoration of the *CAT* gene (T.Q.T. and R.R.S., manuscript in preparation). Here we examine the effect of reversing the *CAT* gene on the frequency of deletion of a

A pBR325n1, Leading Strand

a

```
          <-----           <-----
5′ gctcatccggAATTCGTCTGATGCACGaattccgtatggca 3′
3′ cgagtaggccttaaGCAGACTACGTGCTTAAggcataccgt 5′
```

b

```
          <-----           <-----
5′ gctcatccggAATTCGTCTGATGCACGaattccgtatggca 3′
3′                               CTTAAggcataccgt 5′
                                 <-----
```

c

```
                G C A
              T     C
            A         G
          G           a
            T       a
              C   t
                T t
                  Gc
5′ gctcatccggAATTCcgtatggca 3′
                  CTTAAggcataccgt 5′
                  <-----
```

B pBR325n1, Lagging Strand.

a

```
                -----> ----->
5′ gctcatccggAATTCGTCTGATGCACGaattccgtatggca 3′
3′ cgagtaggccttaaGCAGACTACGTGCTTAAggcataccgt 5′
```

b

```
                ----->
5′ gctcatccggAATTC
3′ cgagtaggccttaaGCAGACTACGTGCTTAAggcataccgt 5′
```

c

```
                ----->
5′ gctcatccggAATTC
3′ cgagtaggcCTTAAggcataccgt 5′
              cG
            t T
          t   G
        a     C
      a       A
    G         T
  C           C
    A G A
```

FIG. 2 Potential structural intermediates involved in deletion of a non-palindromic symmetric deletion insert. *A*, Potential DNA secondary structure in the leading strand of pBR325n1 containing a 'symmetric non-palindromic insert'. *a*, Nucleotide sequence of the 17-bp n1 insert and flanking DNA. The n1 insert is in upper-case bold characters. The direct repeats are represented by arrows above the sequence. *b*, DNA replication of the first copy of the direct repeat 5'GAATTC3'. *c*, The template strand slips and the second (left-most) copy of the direct repeat in the template pairs with the direct repeat in the progeny strand. Replication continues resulting in deletion of the fragment between direct repeats in the progeny strand. *B*, Potential DNA secondary structure in the lagging strand of pBR325n1. Events shown are analogous to that described above for the leading strand.

symmetric (n1) and asymmetric (p1) insert (Fig. 1). These pBR325-based plasmids reverted to a Cm^r phenotype at frequencies of $2\text{-}44 \times 10^{-9}$ Cm^r revertants per viable cell in the *recA*$^-$ *E. coli* strain DH5 (Table 1). Reversing the *CAT* gene in the pBR523-based plasmid did not alter the Cm^r reversion frequency of the symmetric n1 deletion insert. The Cm^r reversion frequencies of two other symmetric deletion inserts (one non-palindromic, n2, and one palindromic, p1) were also similar when the gene was reversed (T.Q.T. and R.R.S., manuscript in preparation). When the *CAT* gene was reversed, the reversion frequency for the asymmetric p1 deletion insert was reduced by about a factor of 20. This suggests that deletion occurs preferentially in either the leading or lagging strand.

The deletion inserts in the *CAT* gene are symmetric or asymmetric with respect to replication (Fig. 1) and the potential for misalignment that can lead to deletion (Figs 2, 3). The n1 deletion insert represents a symmetric situation because mispaired intermediates containing an unpaired loop can form during replication in the leading and lagging strands (Fig. 2). The p1 insert is asymmetric with respect to replication of the direct-repeated and inverted-repeated elements (Fig. 1). As a consequence of this asymmetry, different DNA secondary structure may be involved in deletion in the leading and lagging strand as explained in Fig. 3. In the leading strand of pBR325p misalignment may involve the formation of a loop in the DN between the direct repeats (*A*). In the lagging strand, a hairp may form, stabilizing the misalignment of direct repeats an leading to deletion of the insert (*B*). The secondary structure which may occur at the 3′ end of the leading strand of pBR325p are those that could occur at the 3′ end of the lagging strand pBR523p1. Similarly, events that could occur at the 3′ end the lagging strand of pBR325p1 are those that could also occ at the 3′ end of the leading strand of pBR523p1.

Reversing the *CAT* gene in pBR325p1 decreased the reversion frequency by a factor of 20 (Table 1). The same res has been obtained in two other *recA*$^-$ strains HB101 and JT This result is consistent with the frequency of mutagene differing between the leading and the lagging strands. In lagging strand of pBR325p1, where misalignment may be stab ized by a hairpin, a deletion frequency of 44×10^{-9} was obser In cases of n1 and the lagging strand of pBR523p1, wh deletion may involve the formation of a loop, the revers

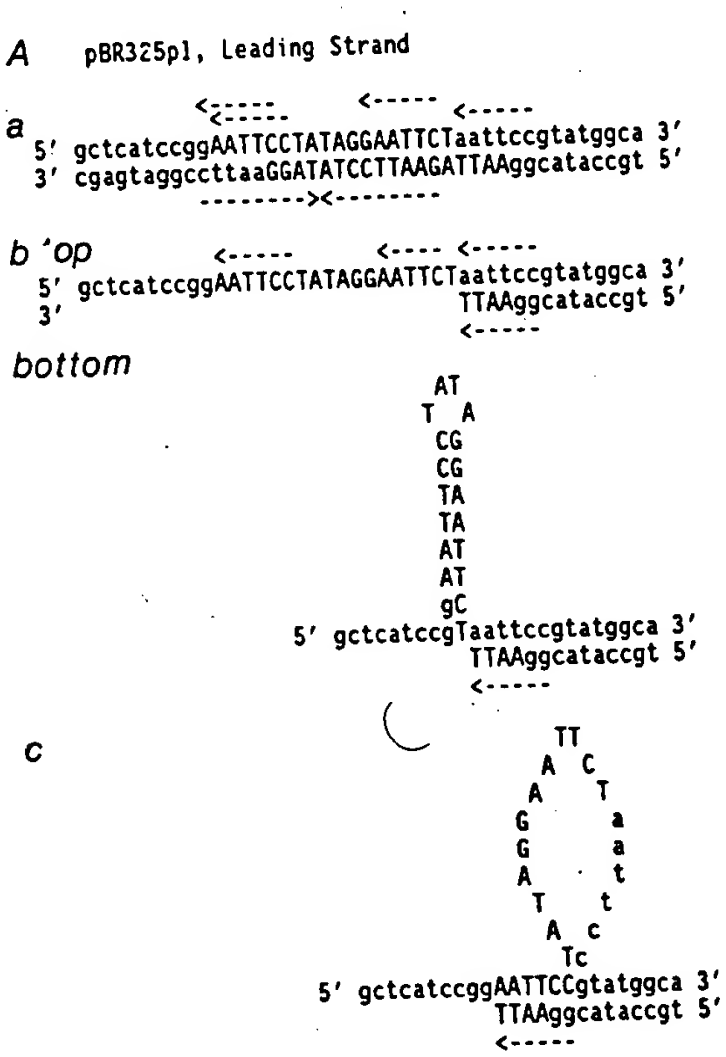


FIG. 3 Potential structural intermediates involved in deletion of a palindromic asymmetric deletion insert. *A*, Potential DNA secondary structure in the leading strand of pBR325p1 containing an 'asymmetric palindromic insert'. *a*, Nucleotide sequence of the 18-bp p1 insert and flanking DNA. *b*, The direct repeats are represented by arrows above the sequence and the inverted repeat is indicated by facing arrows below. *b*, top, DNA replication of the first copy of the direct repeat 5′AATTCC3′ in single stranded DNA. *b*, bottom, Replication in DNA in which the inverted repeat has formed a hairpin arm. Such a structure may temporarily stop the replication fork[15]. Note that the second copy of the direct repeat is now base-paired within the hairpin arm. *c*, The template strand slips to allow mispairing between the second (left most) copy of the direct repeat on the template strand and the first copy of the direct repeat on the progeny strand. This structure may form from either intermediate shown in *b* (top and bottom). This misalignment event requires formation of a loop of DNA that can not form a hairpin arm. *B*, Potential DNA secondary structure in the lagging strand of pBR325p1. *a*, Orientation of direct and inverted repeats. The arrows indicating the direct repe are shown in the direction of replication. *b*, Replication of the first copy of the dir repeat. If the inverted repeat existed as a hairpin before replication, the replicat fork would necessarily begin replication through the hairpin arm to copy the first dir repeat. *c*, Misalignment of the first copy of the direct repeat in the progeny str with the second (right-most) copy of the direct repeat in the template strand. In case, because of the location of the inverted repeat, the misaligned interme contains a perfect hairpin arm in the template strand which may stabilize the m ment. Hairpins in cruciform structures represent Holliday recombinational interm in which, under the proper ionic conditions, there are no unpaired bases[16]. Thu junctions have been shown to exist as stable structures and probably contain unst helices in a 'Y' structure[17]. The conformation of the DNA may well be influe the interaction with DNA polymerase. The structure as drawn would contain mismatch at the three-way junction.

TABLE 1 The frequency of Cm^r reversion mutations

	pBR325, natural orientation		pBR523, *CAT* gene reversed	
	Plasmid	Reversion frequency	Plasmid	Reversion frequency
p1	pBR325p1	$44\times10^{-9}\pm 20\times10^{-9}$	pBR523p1	$2.0\times10^{-9}\pm 0.9\times10^{-9}$
n1	pBR325n1	$3.7\times10^{-9}\pm 0.9\times10^{-9}$	pBR523n1	$2.7\times10^{-9}\pm 0.9\times10^{-9}$

To determine reversion frequencies, 25 ml of *E. coli* DH5 containing the plasmids were grown to a density of 5×10^8 cells ml^{-1} in K medium containing 10 μg ml^{-1} tetracycline from 1 ml of an overnight culture. Reversion frequencies are the number of Cm^r cells in the population, which is determined by plating a known number of cells on Luria plates containing tetracycline and chloramphenicol (25 μg ml^{-1} each). Total viable cells were determined by plating on medium containing tetracycline. The frequency represents the average of 6–8 to individual experiments in which the variation between experiments was typically within a factor of two. Analysis of plasmid copy number showed no detectable difference for any of these plasmids. The structure of the revertants was determined by digestion with *Eco*RI and analysis of the 129-bp *Alu*I fragment, which contained the *Eco*RI site, to establish the size of the deletion to within about 2 bp. The DNA sequence of the region was determined for several revertants from each deletion insert. In >95% of the revertants examined for pBR325p1 (and n2 and p4 inserts) complete deletion of the insert and one copy of the direct repeat occurred. For the n1 plasmids and pBR523p1 the revertants consisted of complete deletions in addition to some revertants which have not completely deleted the insert. We examined 12–24 independent revertants for each plasmid.

frequency was about 10–20-fold lower, or $2\text{–}3.7\times10^{-9}$, consistent with the interpretation that misalignment stabilized by a hairpin has a higher reversion frequency than misalignment in which a loop is formed. Analysis of other inserts provides additional evidence in support of this interpretation (T.Q.T. and R.R.S., manuscript in preparation). Based on these results, if deletion between direct repeats occurred preferentially on the lagging strand, pBR325p1 should have a higher reversion frequency than pBR523p1. The results suggest that misalignment leading to deletion may occur preferentially in the lagging strand and its template.

Although our results are consistent with differential mutagenesis in the leading and lagging strands, we cannot exclude the possibility that differences in reversion frequencies are due to differences in the effects of orientation of these sequences on homologous or illegitimate recombination, or aberrant repair of spontaneous damage. As reversions frequencies were measured in *rec*A$^-$ strains homologous recombination should be minimal and there seems no reason to expect different rates of recombination in the insert when the whole gene is reversed. In addition, if such differences did exist, they might be expected to be seen with all deletion inserts. The correlation of deletion frequencies with potential for the formation and stability of possible replication intermediates therefore strongly supports the hypothesis that the observed deletion events are the result of replication errors (T.Q.T and R.R.S., manuscript in preparation).

Eukaryotic cells repair DNA damage in active genes more efficiently than in inactive genes[10]. In addition, a preference for repair of the transcribed strand has been demonstrated in both *E. coli*[11] and eukaryotic cells[12]. As asymmetric DNA replication complexes are involved in eukaryotes[13] and prokaryotes[6] a difference in the rate of spontaneous mutagenesis in different strands might not be unexpected. An initial investigation did not reveal a large difference in the fidelity of misincorporation during DNA replication in an *in vitro* HeLa cells system[14]. Our results present the first evidence for differential mutagenesis in the leading or lagging strand of *E. coli*. □

Received 12 March; accepted 10 June 1991.

1. Glickman, B. W. & Ripley, L. S. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4046–4050 (1984).
2. Ripley, L. S. & Glickman, B. W. in *Cellular Responses to DNA Damage, UCLA Symposia on Molecular and Cellular Biology*, Vol 10 (ed. N. R. Cozzarelli) 521–540 (Liss, Inc. New York, N.Y. 1983).
3. Drake, J. W., Glickman, B. W. & Ripley, L. S. *Am. Scient.* **71**, 621–630 (1983).
4. Schaaper, R. M., Danforth, B. N. & Glickman, B. W. *J. molec. Biol.* **189**, 273–284 (1986).
5. Ripley, L. S. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4128–4132 (1982).
6. Streisinger, G. *et al. Cold Spring Harb. Symp. quant. Biol.* **31**, 77–84 (1966).
7. Albertini, A. M., Hofer, M. Calos, M. P. & Miller, J. H. *Cell* **29**, 319–328 (1982).
8. McHenry, C. S. *A. Rev. Biochem.* **57**, 519–550 (1988).
9. Scott, J. R. *Microbiol. Rev.* **48**, 1–23 (1984).
10. Bohr, V. A., Smith, C. A., Okumoto, D. A. & Hanawalt, P. C. *Cell* **40**, 359–369 (1985).
11. Mellon, I. & Hanawalt, P. C. *Nature* **342**, 95–98 (1989).
12. Mellon, I., Spivak, G. & Hanawalt, P. C. *Cell* **51**, 241–249 (1988).
13. Tsurimoto, T. Melendy, T. & Stillman, B. *Nature* **346**, 534–539 (1990).
14. Roberts, J. D., Thomas, D. C. & Kunkel, T. A. *Proc. natn. Acad. Sci. U.S.A.* **88**, 3465–3469 (1991).
15. Weaver, D. T. & DePamphilis, M. L. *J. molec. Biol.* **180**, 961–986 (1984).
16. Duckett, D. R., Murchie, A. I. H. & Lilley, D. M. J. *EMBO J.* **9**, 583–590 (1990).
17. Duckett, D. R. & Lilley, D. M. J. *EMBO J* **9**, 1659–1664 (1990).

ACKNOWLEDGEMENTS. This work was supported by grants from the American Cancer Society and the National Institutes of Environmental Health Sciences. We thank K. Allen and E. Schwab for assistance.

ERRATUM

Homozygous prion protein genotype predisposes to sporadic Creutzfeldt–Jakob disease

Mark S. Palmer, Aidan J. Dryden, J. Trevor Hughes & John Collinge

Nature **352**, 340–342 (1991)

IN this letter in the 25 July issue, the sentence beginning in line 1 on page 341 should read: "Of the 22 CJD cases, 16 were homozygous for Met 129, 5 homozygous for Val 129 and only one was a heterozygote, while of the suspected CJD cases, 13 were Met-129 homozygotes, 4 were heterozygotes and 6 were Val-129 homozygotes." The published version incorrectly referred to 11 Met-129 homozygotes.

CORRECTION

Novel myosin heavy chain encoded by murine *dilute* coat colour locus

John A. Mercer, Peter K. Seperack, Marjorie C. Strobel, Neal G. Copeland & Nancy A. Jenkins

Nature **349**, 709–713 (1991)

WE have discovered frameshift errors made in the assembly of our published cDNA sequence affecting only the region of nucleotides 1138–1180, corresponding to amino acid residues 367–379. The correct sequence is available on the EMBL database under accession number X57377, release 28. These errors do not affect the conclusions of the paper in any way. We thank D. Cheney, E. Espreafico, R. Larson, and M. Mooseker for sharing unpublished sequence data.

ELIMINATION OF THE SECONDARY STRUCTURE EFFECT IN GEL SEQUENCING OF NUCLEIC ACIDS

N. S. AMBARTSUMYAN and A. M. MAZO

Institute of Molecular Biology, USSR Academy of Sciences, Moscow 117984, USSR

Received 10 March 1980

1. Introduction

Two basic approaches are now used in sequencing nucleic acids: (1) the introduction of an end label into the molecule and subsequent specific, chemical or enzymatic digestion of polynucleotides [1–4]; (2) that based on enzymatic synthesis using DNA and RNA polymerases [5–9]. In either case, the products are separated by thin-layer electrophoresis in polyacrylamide gels with 7 M urea [1,10]. The disadvantage that the two methods have in common is abnormal mobility of certain oligonucleotides in electrophoresis (compression) which is usually attributed to their stable secondary structure. These structured regions may be melted due to the high temperature of sequencing gel [4].

Here we have modified cytosine residues in DNA by a mixture of methoxyamine and bisulfite. This treatment destroys the formation of GC base pairs and thus eliminates the effect of the secondary structure on the mobility of DNA fragments in gels. We have found that if cDNA synthesized on a template of phage MS2 RNA using reverse transcriptase in the presence of dideoxynucleoside triphosphates is subjected to such a treatment, further electrophoretic separation is considerably improved and certain ambiguities in the sequence are clarified.

This strategy in eliminating the 'compression' has been used earlier in RNA sequencing [11]. Therefore, chemical modification that blocks the formation of secondary structure is of universal significance in gel-sequencing determination of the primary structure of nucleic acids.

2. Materials and methods

2.1. *Materials*

Phage MS2 RNA was a generous gift of Dr V. Berzin' (Institute of Organic Synthesis, Latvian SSR Academy of Sciences). The oligonucleotide $^{5'}$CTCATGTT$^{3'}$ complementary to an intercistronic region between the genes for coat protein and for RNA replicase of phage MS2 was a kind gift from N. F. Sergeyeva and V. Veiko (Moscow State University). Protein S1 from the small subunit of *Escherichia coli* ribosomes was prepared in the laboratory of Professor S. E. Bresler (P. Konstantinov Institute of Nuclear Physics, USSR Academy of Sciences). RNA-dependent DNA polymerase was a generous gift of Professor J. Beard (Life Sciences, USA). Deoxynucleoside triphosphates (dNTP) and dideoxynucleotide triphosphates (ddNTP) were purchased from P-L Biochemicals, [α-^{32}P]dATP, dCTP and dGTP (400 Ci/mmol), from the Radiochemical Centre (Amersham). Reagents for electrophoresis were from Reanal (Hungary). All other reagents used in experiments were of the highest purity grade.

2.2. *DNA synthesis in the presence of chain terminators*

The reaction was conducted in two variants:

(1) The incubation mixture (10 μl) contained 50 mM Tris–HCl (pH 8.3), 6 mM $MgCl_2$, 4 mM dithiotreitol, 400 μg MS2 RNA/ml, a 40-fold molar excess of the primer, 5 μM dNTP, and 1 μCi of one of the labeled triphosphates. ddNTP were added to the incubation mixture at the following

Elsevier/North-Holland Biomedical Press

June 1980
2) Nature New Biol.
tawiski, M. and Harrell,
. Res. 4, 117–162.
'4) Cancer Res. 34,
. and Boutwell, R. K.
569.
980) Cancer Res. 40,
980) Nature 284,
72) Biochim. Biophys.
1 Murray, A. E. (1976)
78) Front. Matrix Biol.
C. and Boutwell, R. K.
–2040.
C. and Boutwell, R. K.
–1670.
De Vellis, J., Inglish, D.
(1978) Adv. Cyclic Nucl.
113, 206–210.
na, M. (1979) FEBS Lett.
and Kownatzki, E. (1980)
., Verma, A. K., Gleason,
ell, R. K. (1979) Cancer

molar ratios: ddTTP:dTTP, 2:1; ddATP:dATP, 4:1; ddGTP:dGTP, 1:1; ddCTP:dCTP, 2:1. The reaction was started by adding the enzyme (3 units of activity), and was performed for 30 min at 46°C, then the concentration of 3 dNTP was raised to 100 μM, the dNTP for which ddNTP was added to 25 μM and ddNTP to the above ratio, and 1 unit of activity of the enzyme was added. The incubation was carried on for a further 1.5 h at 46°C.

(2) The incubation mixture (10 μl) was the same with the exception of ddNTP. The incubation was for 12 min at 46°C with the following addition of a mixture containing 100 μM of the 3 dNTP 5 μM of the triphosphate for which the terminator was introduced, and ddNTP at the following molar ratios: ddATP:dATP, 10:1; ddTTP:dTTP, 5:1; ddCTP:dCTP, 2.5:1; ddGTP:dGTP, 5:1. After incubation for 60 min at 46°C, 1 unit of activity of the enzyme was added and the total mixture was incubated for a further 60 min. Once the synthesis was over, the samples were placed in ice, and the material was precipitated with ethanol adding yeast carrier tRNA

2.3. *Modification with a mixture of methoxyamine and bisulfite*

The dried samples were dissolved in 10 μl of a solution of 1.5 M methoxylamine, 1.5 M KCl (pH 5.0) and 1 M $Na_2S_2O_5$ and kept for 4 h at 37°C [11]. The samples were diluted with water to 1 ml, 20 μg carrier tRNA was added and the material was precipitated by adding 1:10 (v/v) cetyltrimethylammonium bromide (10 mg/ml) in 1 M Na-phosphate buffer (pH 5.0). After centrifugation at 6000 rev./min for 15 min the pellet was dissolved in 2 M NaCl and precipitated with ethanol; the precipitate was dissolved in 0.2 M sodium acetate (pH 5.0) and again precipitated with ethanol.

2.4. *Preparation of samples for electrophoresis*

The precipitates were dissolved in 5 μl H_2O, and 5 μl formamide with 0.01% xylencyanole was added. The mixture was heated at 100°C for 1 min, cooled in an ice bath and then applied to the gel. Electrophoresis in 10% polyacrylamide gel with 7 M urea was performed as in [1]. The dimensions of the gel were 60 × 20 × 0.06 cm. After electrophoresis, the gel was covered with Saran-Wrap and exposed with RT-1 X-ray film at −70°C.

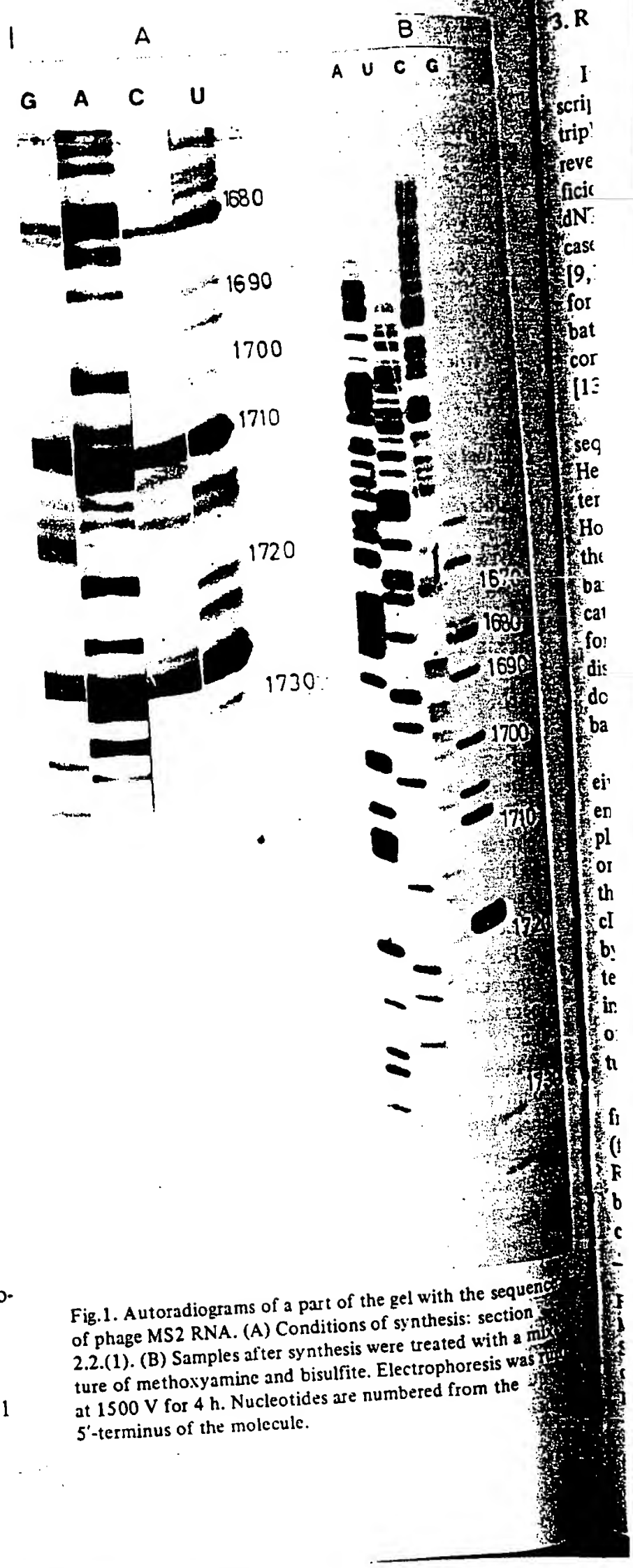
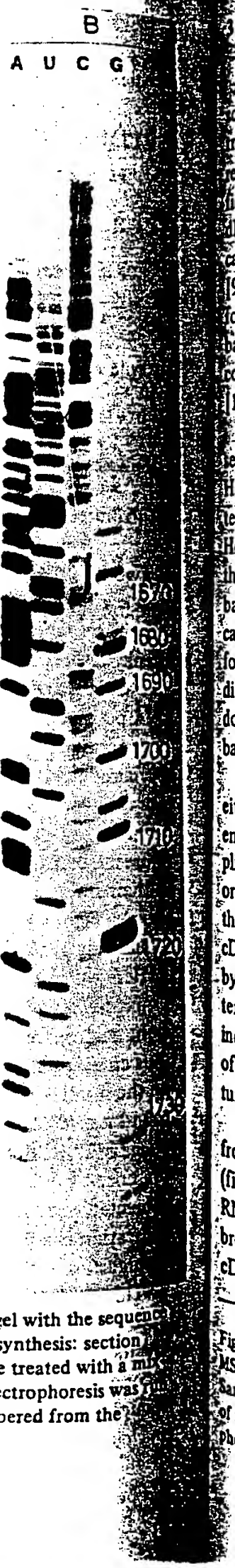


Fig.1. Autoradiograms of a part of the gel with the sequence of phage MS2 RNA. (A) Conditions of synthesis: section 2.2.(1). (B) Samples after synthesis were treated with a mixture of methoxyamine and bisulfite. Electrophoresis was run at 1500 V for 4 h. Nucleotides are numbered from the 5'-terminus of the molecule.



of the gel with the sequenc
ons of synthesis: section
sis were treated with a mi
fite. Electrophoresis was
e numbered from the

3. Results and discussion

In contrast to DNA polymerase, reverse transcriptase has higher affinity for dideoxynucleoside triphosphates (terminators) [12]. Therefore, for reverse transcriptase to synthesize effectively sufficiently long products (cDNA), the ddNTP: dNTP ratio varies from 0.25:1–10:1, whereas in the case of DNA polymerase, the ratio is 100:1 [9,13–16]. Moreover, this ratio must be determined for each terminator. If work is done with various batches of terminators and different templates, the conditions of the synthesis must also be changed [13–16].

Fig.1A shows part of the gel with a known sequence, A_{1675}–G_{1738}, of RNA phage MS2 [17]. Here, we preliminarily selected the ratio between the terminator and the dNTP for each of the 4 pairs. However, though we made an attempt to optimize the conditions, one can see that the intensity of the bands varies among different tracks, thus complicating the sequencing. Furthermore, bands common for all the terminators (1728, 1710, 1680) can be discerned in several positions in the gel; these bands do not make it possible to establish the nature of bases in these positions of the sequence.

We presumed that such common bands appeared either as the result of stable secondary structure present in these regions on the RNA template (thus complicating their transcription by reverse transcriptase), or originated from 'compression' in the gel caused by the RNA template incompletely separated from cDNA. However, attempts to get rid of these bands by either using alkaline hydrolysis of the RNA template (0.3 M KOH, 37°C, 16 h) or adding to the incubation mixture protein S1 from the small subunit of *E. coli* ribosomes that melts the secondary structure of RNA was unsuccessful (not shown).

These common bands almost entirely disappear from all columns if fresh RNA preparations are used (fig. 2A). We believe that specific breaks appear in the RNA molecule in the course of its storage; these breaks account for significant 'structural stops' in cDNA synthesis.

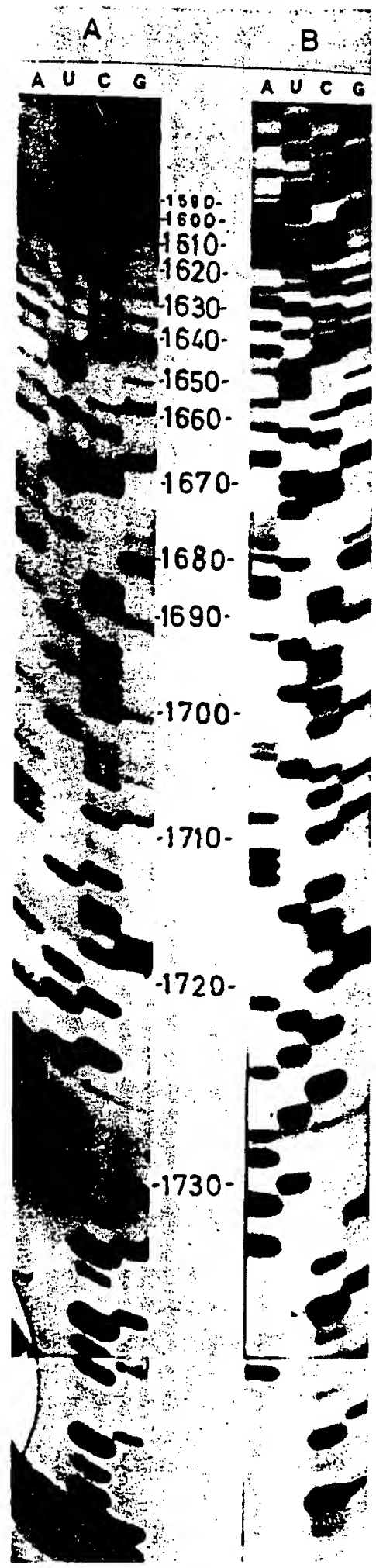


Fig.2. Autoradiograms of the gel with the sequence of phage MS2 RNA. (A) Conditions of synthesis: section 2.2.(2). (B) Samples after cDNA synthesis were treated with a mixture of methoxyamine and bisulfite. Conditions of the electrophoresis and numbering of nucleotides were as in fig.1.

As can be seen in fig.1B, if terminators are added at the beginning of the reaction (variant (1), section 2.2), differences in the length of cDNAs produced are observed, besides differences in the intensity of bands in different tracks. In order to eliminate these differences, we changed the conditions of cDNA synthesis (fig.2A). In this case, terminators were added 12 min after the beginning of the synthesis (variant 2, section 2.2). Under these conditions, the difference in the intensity of bands was considerably diminished, as well as that in the length of the products. The effectiveness of termination however did not change. The addition of terminators (with the increase in their relative content) sharply decelerated the synthesis; therefore the length of the products was predetermined mainly at the first stage of the reaction (without ddNTP) if terminators were added after the beginning of the reaction.

Fig.2A shows that though considerable improvement has been achieved as compared to fig.1, the sequence can be read for a relatively short region and, moreover, some regions remain ambiguous (1703–1708, 1718, and upstream from 1670).

In order to improve separation in the gel, we modified cDNA with a mixture of methoxyamine and bisulfite which resulted in the selective and quantitative conversion of cytidine into 5,6-dihydro-6-sulfo-N^4-methoxycytidine [18]. As can be seen in fig.1B, such a modification significantly improves the separation of polynucleotides. The sequence is read easily and unambiguously; moreover, the sequence being read is much longer in fig.2B than in fig.1B, 2A, and reaches 150 nucleotides.

As we have shown, the modification makes the secondary structure of tRNA and 5 S RNA unfold considerably, which may be attributed apparently to weakening of GC base pairs. As a result, 'compressions' caused by the secondary structure of RNA disappear from the sequencing gel.

Therefore, the strategy for elimination of defects in electrophoretic separation of polyribonucleotides in [19] is applicable, as has been shown here, for DNA and may be of universal significance in gel-sequencing of nucleic acids.

The sequence of the region 1710–1580 slightly differs from that determined in [17]. According to our data, positions A_{1699} and G_{1663} [17] are occupied by cytosines. Presumably, these differences may be attributed to strain differences in phage MS2.

When this manuscript was in preparation, we discovered a publication [19] in which 'compressio' of the products of Qβ replicase synthesis was elimi nated by adding ITP rather than GTP as a precurso to the incubation mixture.

Acknowledgements

The authors are greatly indebted to Professor I. Kisselev for his stimulating discussions and continu support in this work, to Professor J. Beard for a generous gift of reverse transcriptase, to V. Berzin N. Sergeyeva and V. Veiko for kind gifts of the phage MS2 RNA preparations and the oligonucleo ^{5}'CTCATGTT^3'.

References

[1] Maxam, A. M. and Gilbert, W. (1977) Proc. Natl. A Sci. USA 74, 560–564.
[2] Donis-Keller, H., Maxam, A. M. and Gilbert, W. (19 Nucleic Acids Res. 4, 2527–2538.
[3] Simoncsits, A., Brownlee, G. G., Brown, R. S., Rub J. R. and Guilley, H. (1977) Nature 269, 833–836
[4] Peattie, D. A. (1979) Proc. Natl. Acad. Sci. USA 76 1760–1764.
[5] Sanger, F. and Coulson, A. R. (1975) J. Mol. Biol. 94 4441–4448.
[6] Brownlee, G. G. and Cartwright, E. M. (1977) J. Mo Biol. 114, 93–117.
[7] Kramer, F. R. and Mills, D. R. (1978) Proc. Natl. Ac Sci. USA 75, 5334–5338.
[8] Axelrod, V. D., Vartikyan, R. M., Aivazashvili, V. A. and Beabealashvili, R. Sh. (1978) Nucleic Acids Re 3549–3563.
[9] Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Pro Natl. Acad. Sci. USA 76, 731–735.
[10] Sanger, F. and Coulson, A. R. (1978) FEBS Lett. 87 107–110.
[11] Mazo, A. M., Mashkova, T. D., Avdonina, T. A., Ambartsumyan, N. S. and Kisselev, L. L. (1979) Nucleic Acids Res. 7, 2469–2482.
[12] Smoler, D., Molineux, I. and Baltimore, D. (1971) J. Biol. Chem. 246, 7697–7700.
[13] McGeoch, D. J. and Turnbull, M. T. (1978) Nucleic Acids Res. 6, 4007–4024.
[14] McGeoch, D. J. and Dolan, A. (1979) Nucleic Acids Re 6, 3199–3211.
[15] Zimmern, D. and Kaesberg, P. (1978) Proc. Natl. A Sci. USA 75, 4257–4261.
[16] Bina-Stein, M., Thoren, M., Salzman, N. and Thomp J. A. (1979) Proc. Natl. Acad. Sci. USA 76, 731–7
[17] Fiers, W., Contreras, R., Duernick, F., Haegeman Iserentant, D., Merregaert, J., Min Jou, W., Molem F., Raeymaekers, A., Van den Berghe, A., Volcka and Ysebaert, M. (1976) Nature 260, 500–507
[18] Sverdlov, E. D., Monastyrskaya, G. S., Gushkov Levitan, T. L., Scheichenko, V. I. and Budovsky (1974) Biochim. Biophys. Acta 340, 153–165
[19] Mills, O. and Kramer, F. R. (1979) Proc. Natl. USA 76, 2232–2235.